Growing a Bayesian Conspiracy Theorist: An Agent-Based Model

Jens Koed Madsen (jens.madsen@ouce.ox.ac.uk)
School of Geography and the Environment, University of Oxford
OX1 3QY, South Parks Road, Oxford, United Kingdom
Orcid: 0000-0003-2405-8496

Richard Bailey (richard.bailey@ouce.ox.ac.uk)
School of Geography and the Environment, University of Oxford
OX1 3QY, South Parks Road, Oxford, United Kingdom

Toby D. Pilditch (t.pilditch@ucl.ac.uk)
Department of Experimental Psychology, University College London
WC1E 6BT, Gower Street, London, United Kingdom

Abstract

Conspiracy theories cover topics from politicians to world events. Frequently, proponents of conspiracies hold these beliefs strongly despite available evidence that may challenge or disprove them. Therefore, conspiratorial reasoning has often been described as illegitimate or flawed. In the paper, we explore the possibility of growing a rational (Bayesian) conspiracy theorist through an Agent-Based Model. The agent has reasonable constraints on access to the total information as well its access to the global population.

The model shows that network structures are central to maintain objectively mistaken beliefs. Increasing the size of the available network yielded increased confidence in mistaken beliefs and subsequent network pruning, allowing for belief purism. Rather than ameliorating and correcting mistaken beliefs (where agents move toward the correct mean), large networks appear to maintain and strengthen them. As such, large networks may increase the potential for belief polarization, extreme beliefs, and conspiratorial thinking – even amongst Bayesian agents.

Keywords: Conspiratorial thinking; Extreme beliefs; Agent-Based Models; Bayesian updating

Introduction

Truth is the shattered mirror strewn in myriad bits; while each believe his little bit the whole to own
Richard Burton (The Kasîdah of Hâjî Abdû El-Yezdî)

In recent years, scientists, scholars, and commentators have remarked upon the apparent rise of epistemic echo chambers (see e.g., Bakshy et al., 2016) and increasing political polarization. Echo chambers refer to communities, online or otherwise, that interact predominantly with themselves and who rarely, if ever, seek information aside from the information available within the chamber. Whether endogenously created (such as cults) or exogenously created (such as living on an island with no contact to the outside world), the emergence and maintenance of epistemically enclosed systems and their consequences is interesting and worth studying. The current paper explores the possibility of generating, maintaining and strengthening encapsulated belief communities through an Agent-Based Model (Gilbert, 2008) where every agent is rational (here, Bayesian) and where information is potentially available to challenge the viewpoint of the agent.

Specifically, we are interested in exploring the possibility of generating conspiratorial beliefs. That is, beliefs that are maintained despite being objectively false and there being available evidence to challenge or disprove the theory in question. Proponents of such beliefs frequently hold these positions strongly. We explore whether it is possible to strengthen confidence in objectively mistaken beliefs through a rational process given imperfect knowledge about the world. Rather than assuming illegitimate updating processes or special cognitive functionality, the model tests if, in principle, a Bayesian conspiracy theorist can emerge and be maintained. That is, the model explores whether or not individual differences are a necessary requirement for the emergence and maintenance of extreme beliefs.

The Burton quote at the top of the introduction can be seen as foundational for the paper. It suggests that beliefs can be generated and maintained as a fragment of a larger, and often very different, picture. Further, it intimates that humans generate inferences about the world on the back of the evidence available to us at any given time in our lives. This information may come through first-hand experience or through other sources such as parents, peers, media outlets, and experts.

In order to set the scene for the Agent-Based Model, we briefly consider how conspiratorial thinking has previously been approached in the literature.

Conspiratorial thinking

Conspiracy theories can be loosely defined as beliefs that are held strongly when evidence is broadly available to challenge or entirely refute the theory. Yet, proponents maintain (and might even strengthen) their belief in the theory despite the availability of this evidence. However, in order to adequately simulate emerging conspiracy theories, we need to employ a more stringent definition of conspiratorial thinking.

According to Barkun (2003), conspiracy theories are characterized by three traits. First, conspiracy theories operate under the assumption that nothing happens by accident. From a cognitive perspective, this may be described as causal oversensitivity where the reasoner generates causal links between disparate and supposedly separate pieces of information, leading to over-connection. Second, Barkun argues that for conspiracy theorists, nothing is really what it appears to be on the surface (i.e. the 'real' causal mechanisms between pieces of information is covered up(?) by any official story). This element, too, suggests a cognitive agent who over-weights and over-generates causal links between independent pieces of information. For example, some proponents of the moon landing conspiracy theory believes the director of *A Space Oddysey: 2001*, Stanley Kubrick, to be involved in producing faked photography because Kubrick hired crew for 2001 who used to work for NASA (Frederick Orway and harry Lange). Finally, Barkun argues that conspiracy theorists tend to believe things to be highly connected. To this end, Barkun argues that conspiracy theories eventually become enclosed systems that are falsifiable if confronted with additional evidence and therefore "a matter of faith rather than proof". Presumably, this entails that conspiracy theorists might stop seeking new information and instead assume their beliefs to be a priori true. As evident from these definitions, the operationalizing of conspiracy theories usually involves special cognitive make-up and a heuristic process that treat information related to the conspiracy theory as qualitatively different from 'normal' belief updating. The current model explores whether these are valid assumptions.

Indeed, Birchall (2006) describes conspiratorial thinking as illegitimate updating and belief maintenance (as opposed to normative, legitimate reasoning). In general, conspiratorial thinking is typically conceived as an abnormal and potentially fallacious (or illegitimate) reasoning process, which relies heavily on cognitively biased heuristics such as over-generation of causal links, erroneous attribution of motives, and mistaken perception of interconnectivity. Commonly, these conspiratorial thinking accounts assume conspiracies are a product of mistaken or misguided reasoning.

In this paper, we provide a proof of concept that conspiratorial thinking can emerge from Bayesian rational paradigms given access to a subset of evidence and the possibility of interacting with other like-minded agents. As will be argued later, we believe both of these assumptions to be realistic and grounded in psychological findings. As will be explained further in the paper, we show that conspiratorial agents do not require special cognitive abilities or predispositions in order to be supremely confident in their (objectively mistaken) belief. This approach is reminiscent of work conceptualizing supposed reasoning flaws through cognitively reasonable processes. This work includes, but is not limited to, Bayesian accounts of argument fallacies (e.g., Corner et al., 2011; Harris et al., 2012), a Bayesian model of appeals to authority (e.g. Hahn et al., 209; Harris et al., 2015), and skepticism in climate change (Cook & Lewandoswky, 2016). Further, Bayesian agents represent a rational process of integrating new information with prior beliefs pertaining to that hypthesis. For this reason, Bayesian agents have been used previously to explore belief diffusion in networks (see e.g. Jern et al., 2009; Olsson, 2013; Denrell & Le Mens, 2017).

While the current work builds on similar Bayesian accounts of belief updating, we provide a novel contribution to the field by implementing a computational, Agent-Based Model that allows for interaction between agents across time.

For the purpose of this paper, we take conspiratorial thinking to be strongly held beliefs that depart from the objective mean where evidence is available to challenge or refute the theory. Barkun and Birchall argue that these beliefs arise from mistaken or flawed heuristics and/or illegitimate reasoning processes that bias proponents of conspiracy theories toward connectivity, attribution of hidden intentions and over-generation of causal structures. Further, Grimes (2016) argues that conspiratorial beliefs are untenable with larger network structures, as the available information to challenge erroneous views increases. As discussed below, there are some potential challenges for conspiracy accounts that assume special cognitive functions such as oversensitivity toward causal connections as the default cognitive foundation.

Challenges for traditional accounts

The traditional perspectives on conspiratorial thinking may be challenged on at least two grounds. First, it is potentially problematic to ascribe different cognitive functions to the emergence of conspiratorial thinking for two reasons. For one, it is unclear whether this type of reasoning would permeate all beliefs held by that individual (e.g. would a conspiracy theorist also be prone to over-generate causal structures in billiards or snooker). If it were not systemic, it would (insufficiently) appear to be a post hoc account of a particular belief that happens to exhibit such properties. For another, it would not represent a process account of conspiratorial thinking. Rather, it would *assume* differences and apply these to arrive at the conclusion. Instead, we explore whether it is possible to generate objectively mistaken beliefs from the same cognitive processes that generate objectively true beliefs. Both of these would remove the expectancy of abnormality on the part of the conspiracy theorist.

Second, traditional accounts tend to focus on the cognitive function of the isolated individual, rather than on systemic belief diffusion as a result of interactions with other people. As discussed in the following, studies and simulations have shown that aggregate behavioral patterns might not be reducible to the components in isolation if the components can interact with each other in meaningful ways (see e.g. Johnson, 2001; Ball, 2005) given complex and dynamic environments (Johnson, 2009). Faced with the problem of epistemic isolation, we apply Agent-Based Modeling to explore the potential of growing a Bayesian conspiracy theorist without adding special cognitive functions to the agent in question.

Agent-Based Modeling

In order to circumvent the problems caused by traditional individual-based accounts of cognition, we employ Agent-Based Modeling, which allows for simulations of belief networks populated by Bayesian agents. This further allows for introduction of heterogeneity, as will be discussed later (here, initial sampling allows agents to gather and evaluate data individually, which provides heterogeneous priors). Agent-Based Models (ABMs) are computer simulated multi-agent systems that describe the behavior of and interactions between individual agents, who operate in synthetic environments (Gilbert, 2008; Bandini et al., 2009). Agents are encoded on a computational basis, and may implement and explore models of cognitive function. They allow for complex, dynamic and adaptive systems to emerge through interactions between agents and with the environment as well as across time (Miller & Page, 2007). ABMs may be described in terms of their three fundamental components: Agents, Links, and Patches.

Agents are the nodes representing the active cognitive entities of the system. They can make decisions and make use of information in any way that is formally expressible. These functions include, but are in no way limited to, utility valuations, Bayesian belief inferences, stock market engagement, and so forth. The agents may reproduce (e.g. give birth to a new agent), move around the simulated space, and make new (and potentially more relevant) decisions as they learn more about the environment. In order to engage with the environment, agents will have specified rules for agent-environment interactions such as fishing, purchasing a house, moving around the simulated space and so forth. These behaviors and inferences may yield dynamic and adaptive aggregate behavioral patterns. For example, if all agents harvest simultaneously, Tragedy of the Commons type problems (Ostrom, 2012) can emerge. In the present model, we allow for Bayesian belief updating as the agents encounter new information or talk with other agents via links.

Links represent rules for possible interactions between agents. Links can be any interactivity that can be expressed formally. The interaction may be direct (e.g. communication between two agents or sales structure between agents, see Epstein & Axtell, 1996) or indirect (e.g. social attraction or repulsion or emotional feedback, see Schelling, 2006; Epstein, 2013). Interactions allow for feedback loops to emerge, which in turn may generate aggregate behavioral patterns that are irreducible to the components in isolation.

Patches represent the simulated environment in which agents exist. They can have any and all properties that are formally describable. If consumable (such as grass for sheep, fish for fishers), they may give the agent energy, money, or other affordances. Patches may be dynamic such that they might regrow or migrate. Further, patches may facilitate or restrict movement of agents in the simulated space. The patches provide the foundational and potentially dynamic environment in which the agents live and act. In the model we present, the environment restricts interaction between agents if the search potential is low.

Compared with traditional methods, ABMs are capable of simulating dynamic and adaptive decision-making in changeable environments (Miller & Page, 2007). This allows for agents to self-organize without hard-wiring expected aggregate behavioral patterns such as emergent echo chambers. Rather, ABMs allow for these properties to emerge, or, in the terminology of Epstein and Axtell (1996), to grow. ABMs further allow for agent and environmental heterogeneity (i.e. agents with different cognitive capabilities).

Growing a Bayesian conspiracy theorist

The aim of the current model is to test a proof of principle that conspiracy theorists can emerge through entirely rational processes without providing any special cognitive functions, heuristic strategies, or access to unique information. In order to do so, we generate an Agent-Based Model where agents can sample information, communicate with one another, and update their beliefs about the world.

Given this initial proof of concept, we simplify the epistemic challenge and consider only one abstract belief. The true probability of the Gaussian distribution from which the agents sample is 0.5. The standard deviation can be manipulated to represent greater or lesser noise in the information environment. In the present paper, the standard deviation is set to 0.2. For the sake of understanding, the probability may represent the belief in the fairness of a coin. If the coin is fair, the distribution of tosses is trivially 50-50 between heads and tails. However, if the coin is not fair, the distribution can be skewed in the direction of either heads or tails. Understood in this way, the agents try to understand if they are in a world in which the coin is fair (uncovering, as it were, the true, underlying probabilities) or if they are in a world where the coin is rigged to either side (arriving at an objectively mistaken belief).

If agents are able to generate, maintain and possibly strengthen a mistaken belief in the epistemic state of the belief, the agent will have exhibited conspiratorial traits, as this fulfills the criterion for the definition in the above: a potentially strongly held, yet objectively mistaken belief, availability of information to challenge or refute the theory, and access to that information. The literature review uncovered two central positions that we explore here. One, we explore Grimes' (2016) argument that conspiratorial thinking is untenable in a large network structure. If this is true, we should see a global regression towards the objectively true mean given larger networks (that is, fewer agents who believe they are in a rigged coin world). Two, we explore Barkun and Birchall's arguments that conspiratorial thinking relies on illegitimate reasoning and biased heuristics. As will be described below, the agents in the model are perfect Bayesian reasoners. If conspiratorial thinking requires special cognitive properties, we should not expect the Bayesian agents to generate strong and mistaken beliefs about the world. The model implements six key elements: generation of prior beliefs, constrained search, network generation, communication between agents, belief updating, and network pruning.

In order to generate a subjective prior belief, agents are born onto the world and sample randomly generated data from a Gaussian distribution (μ = 0.5, σ = 0.2). In a frequentist manner, these are used to calculate a perceived mean and probability density. The sampling represents the worldview of each particular agent *before* they are able to communicate with other agents.

Having generated a prior belief for each agent (and thus introduced sampling heterogeneity), the model relies on four additional assumptions and mechanisms. First, agents cannot sample all available data in the simulated world. This means that they do not have access to all data sampling that other agents have encountered unless they communicate with the other agent in order to learn the beliefs of that agent. As such, agents do not have perfect and complete knowledge about the world in which they live. We believe this is a reasonable assumption, as humans do not have perfect knowledge in real life. Second, although all other agents are hypothetically available, agents cannot communicate to every other agent in the simulated world. Rather, each agent randomly generates the amount of possible communication links. Like the first assumption, we believe this is a reasonable assumption, as humans in the real world cannot communicate with every other person on the planet, but has to settle for a subset of all living persons.

Third, in order to make the agents rational, they update their beliefs about the world in a Bayesian manner. Bayesian updating represents the rational integration of prior beliefs with new evidence to generate posterior belief in the hypothesis. This approach has been applied to a host of related phenomena such as argumentation (Hahn & Oaksford, 2006; 2007), source credibility (Bovens and Hartmann, 2003; Harris et al., 2015), and reasoning and decision-making (Oaksford & Chater, 2007). The integration is formally expressed through Bayes' theorem

$$P(h \mid e) = \frac{P(h)P(e \mid h)}{P(e)}$$

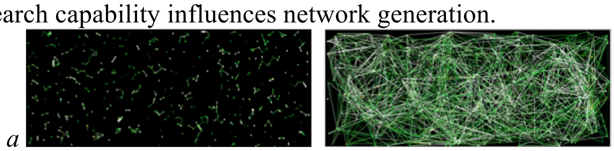
where p(h|e) denotes the posterior belief in the hypothesis (h) given the evidence (e). As such, agents treat each new encounter as a data point to be integrated within their subjective probability density function. Bayesian updating ensures that the agents are fully rational in their belief revision when encountering new evidence.

Finally, several studies on confirmation bias, selectivity bias, and in-group behavior strongly suggest that agents are not entirely stochastic and non-directed in their information search. Taking inspiration from segregation studies (e.g., Schelling, 2006), we introduce a mild preference for people who remotely share their beliefs about the world. The agents are relatively tolerant and will engage in conversation with any other agent who is within ± 1.5 standard deviations of its own perception of the world. Given Gaussian distributions, this means that the agent will speak to 86.6% of people within its belief distribution. Thus, they are willing to talk to and integrate information from agents who have different viewpoints than their own. However, if they are confident in their belief, they will engage with less diverging viewpoints, as their probability density narrows. As an analogy, this means that an agent might be willing to discuss political questions with people with different points of view, but would refuse to engage in discussion with people who believe that fair coin-flips are 60-40 rather than 50-50 in cases where they are absolutely certain about the latter and less certain about the former.

In sum, the agent is born into the world by sampling randomly generated pieces of information related to the hypothesis in question. This informs the mean and standard deviation of their prior. Second, the agents generate networks with other agents within their network radius (which may be limited or encompass the full system). Having set up the model, the agents will communicate freely and honestly (i.e. representing their belief in the hypothesis to the best of their ability), which enables Bayesian belief updating. Agents will maintain communication networks with other agents who are within 1.5 standard deviation of their subjective understanding of the world (i.e. their belief in the hypothesis). If agents within the network fall outside of those boundaries, the agent deactivates the network contact with that particular agent. If agents cannot find any suitable agents within their range, they decrease confidence (simulating negative feedback) and thus expand acceptable search parameters for the following tick. This allows for dynamic network pruning (Ngampruetikorn & Stephens, 2015).

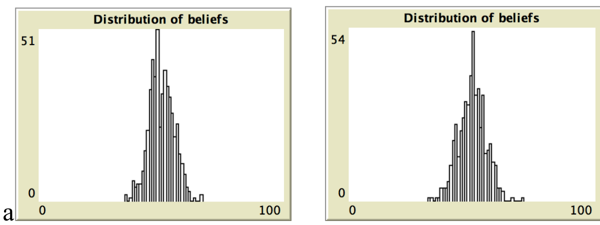
Main findings: Limited and extended networks

We implemented the above model in NetLogo (5.2.1) and manipulated the model in terms of the size of the network. For limited networks, agents had a search range of 10 of 100 (as a product of their geographical location). Extended networks, on the other hand, had a search range of 80 of 100. Agents could connect to and sample randomly from other agents within agent search range who fall within their network criteria. Figs 1a and 1b show the extent to which search capability influences network generation.

a  b

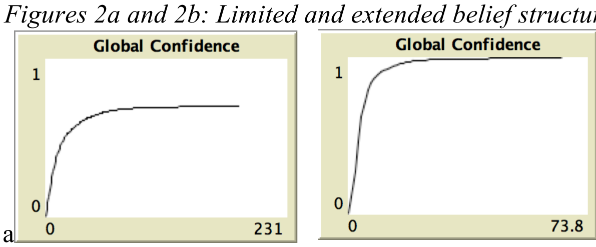
Figures 1a and 1b: Limited and extensive networks

The overall belief structure did not differ significantly between limited and extended networks. Some, but not all agents regressed towards toward the mean while some agents retained their objectively mistaken belief (see Fig. 2a and 2b, which are histograms where number of believers are on the y axis and agent belief is on the x axis), we observe differences in belief confidence. As seen in Figs. 3a and 3b, extended networks allowed for interactions with increasingly like-minded agents, which in turn increased belief confidence. This is true *both* for agents who obtain objectively true and false beliefs. As agents become increasingly confident, their probability density narrows, meaning that they are less willing to engage with agents with differing beliefs. Extended networks allow them to form and maintain contact with agents who share their *specific* beliefs such that they increase their confidence in that particular view of the world. This means purification of beliefs and purification of networks, i.e. the emergence of epistemic echo chambers.



a b

Figures 2a and 2b: Limited and extended belief structures



a b

Figures 3a and 3b: Limited and extended confidence (0-1)

Overall, the model shows that fully rational agents can maintain and potentially strengthen objectively mistaken beliefs. Further, given a mild preference for interaction with like-minded agents, we observe the rise echo chambers. This effect is strengthened with the size of the network. Rather than making extreme beliefs untenable as predicted by Grimes, we show that large networks, here quantified in terms of the number of reachable agents for any given agent, can engender extreme belief maintenance and belief purism.

Discussion and concluding remarks

The Agent-Based Model in the paper provides a theoretical proof of concept that a Bayesian agent can become an ardent conspiracy theorist under three main assumptions. One, the agent does not have perfect and full access to all available information that exists in the world, but can only sample a sub-set of that information. This means that the agent does not rely on perfect knowledge of the system. Depending on the practical conceptualization of information accessibility, the agent may have access to very limited or more extended amounts of information. Two, the agent cannot talk to every other person in the world, but can only talk to a sub-set of all existing agents. Similar to assumption one, this means the agent cannot converse with all other agents and learn their subjective access to information. In the current model, information after prior sampling is gleaned through interactions with other agents. Consequently, by limiting the amount of other agents with whom an agent can engage, the model naturally also limits the access to available information. Principles one and two are concerned with the degree to which the agent can sample information and learn about the world. Three, agents search for and interact with other agents on the basis of their current worldview. They are willing to communicate with most other agents, but avoid other agents with whom they radically disagree about the nature of the world.

The Rise of Echo Chambers

Together, these three (we believe reasonable) assumptions show that larger networks do not yield belief amelioration (as was postulated by some theoreticians who believed the Internet to facilitate greater communication between people and thereby allow a global regression towards the mean). Rather, the model shows that extended networks, given plausible constraints to exposure, lead to the growth of echo chambers and eventual belief purism, whereby agents increasingly discard those who do not share their *specific* beliefs about the world.

One might compare this increasing belief purism to development of political ideologies. In a limited network structure (e.g., a small village), the model suggests that left-leaning voters are willing to communicate with other left-leaning voters (and some right-winged voters depending on the mean and probability density function of the specific voter, mutatis mutandis for right-winged voters). However, in an extended network structure (such as a metropolis or Facebook), the model suggests that voters will have access to other voters who have more similar worldviews. This allows for emergence of political echo chambers where extreme voters have access to other extreme voters. From this, greater belief confidence grows and network pruning increases, as belief purism emerges. We therefore expect increases in network structures will facilitate rather than hinder belief extremism and confidence in worldviews.

The model presented in the current paper allows for this dynamic adaption. In the beginning, agents cluster around people with whom they share general beliefs about the world. However, as they increase in confidence, their probability densities narrow, meaning that fewer agents will fit within the ± 1.5 standard deviations of the perceived mean. As the agent becomes increasingly confident in its own reading of the world, it will be decreasingly inclined to engage with agents who entertain different viewpoints. This allows for belief communities to fracture and radical and supremely confident cells to emerge. The emergent echo chambers function as cyclical maintenance of a peculiar belief.

This finding is interesting because larger networks did not yield belief amelioration, but rather belief solidification. It opens up for a novel way to approach and model epistemic communities that maintain strong beliefs despite available data challenging their beliefs (e.g., creationists, climate skeptics, and radicalized or discriminatory beliefs).

Emergence of reasonably mistaken views

Central to the model, the agents do not have full and perfect knowledge of the world and can only talk to a sub-set of other existing humans. Given the fact that agents update their beliefs in a Bayesian manner, their cognitive system can be described as rational and entirely reasonable. Yet, given incomplete access to data and given the network properties, the model shows that the agents can become entirely confident in objectively mistaken views. As such, we show that extreme beliefs such as conspiracies could emerge through entirely rational processes. While this does not preclude heuristic strategies or special cognitive functions, the model shows that these are not *necessary* for strongly held mistaken beliefs to emerge. Aside from emerging, mistaken beliefs are also able to survive (and even strengthen) in such an environment rather than being swallowed by mainstream beliefs.

Further, agents had a mild preference for communicating with like-minded agents. Rather than making extreme beliefs untenable, the model suggests that increasing the size of the network intensifies the process of radicalization and augments the confidence even in an objectively mistaken belief. In the age of the Internet, this finding is worth considering seriously and exploring further

In conclusion, we have provided a proof of concept that shows the impact of network structures in generating and maintaining extreme beliefs such as conspiratorial thinking. A Bayesian agent can generate and even increase its confidence in objectively mistaken beliefs.

References

Ball, P. (2005) *Critical Mass: How one things leads to another*, Random House, London: UK

Bakhsy, E., Messing, S. & Adamic, L. A. (2016) Exposure to ideologically diverse news and opinion on Facebook, *Science* 348 (6239), 1130-1132

Bandini, S., Manzoni, S. & Vizzari, G. (2009) Agent Based Modeling and Simulation: An Informatics Perspective, *Journal of Artificial Societies and Social Simulation* 12 (4), 1-16

Barkun, M. (2003) *A Cultural of Conspiracy: Apocalyptic Visions in Contemporary America*, University of California Press

Birchall, C. (2006) *Knowledge Goes Pop: From Conspiracy Theory to Gossip*, Berg Publishers, Oxford: UK

Bovens, L., & Hartmann, S. (2003). *Bayesian epistemology.* Oxford: Oxford University Press

Cook, J. & Lewandowsky, S. (2016) Rational irrationality: Modeling climate change belief polarization using Bayesian networks, *Topics in Cognitive Sciences* 8, 160-179

Corner, A., Hahn, U. & Oaksford, M. (2011). The psychological mechanism of the slippery slope argument. *Journal of Memory & Language, 64,* 133-152.

Denrell, J. & Le Mens, G. (2017) Information Sampling, Belief Synchronization, and Collective Illusions, *Management Science* 63 (2), 528-547

Epstein, J. (2013) *Agent_Zero: Toward Neurocognitive foundations For Generative Social Science*, Princeton University Press

Epstein, J. & Axtell, R. (1996) *Growing Artificial Societies*: *Social Science from the Bottom Up*, MIT Press

Gilbert, N. (2008) *Agent-Based Models*, SAGE Publications

Grimes, D. R. (2016) On the Viability of Conspiratorial Beliefs, *PLoS One* 11 (1), e0147905

Hahn, U., Harris, A. J. L., & Corner, A. (2009). Argument content and argument source: An exploration. *Informal Logic, 29,* 337-367.

Hahn, U., & Oaksford, M. (2006) A normative theory of argument strength, *Informal Logic* 26*,* 1-24

Hahn, U., & Oaksford, M. (2007) The rationality of informal argumentation: A Bayesian approach to reasoning fallacies, *Psychological Review 114*, 704-732

Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2015). The Appeal to Expert Opinion: Quantitative support for a Bayesian Network Approach. *Cognitive Science* 40*,* 1496-1533

Harris, A., Hsu, A. & Madsen, J. K. (2012) Because Hitler did it! Quantitative tests of Bayesian argumentation using *Ad Hominem*, *Thinking & Reasoning* 18 (3), 311-343

Jern, A., Chang, K-M. & Kemp, C. (2009) Bayesian belief polarization, *Advances in Neural Information Processing Systems* 22 *(NIPS 2009)*

Johnson, N. (2009) *Simply Complexity: A Clear Guide to Complexity Theory*, Oneworld Publications

Johnson, S. (2001) *Emergence*, Penguin Publications

Miller, J. H. & Page, S. E. (2007) *Complex adaptive systems: An introduction to computional models of social life*, Princeton University Press

Ngampruetikorn, V. & Stephens, G. J. (2015) Bias, Belief, and Consensus: Collective opinion formation on fluctuating networks, *arXiv* 1512.09074v1

Oaksford, M. & Chater, N. (2007) *Bayesian Rationality: The probabilistic approach to human reasoning*. Oxford, UK: Oxford University Press

Olsson, E. J. (2013) A Bayesian simulation model of group deliberation and polarization, in Zenker, F. (Ed.) *Bayesian Argumentation*, Synthese Library, Springer

Ostrom, E. (2012) *The Future of the Commons: Beyond Market Failure and Government regulation*, The Institute of Economic Affairs

Schelling, T. (2006) *Micromotives and Macrobehavior*, Norton and Company, New York: NY