Cascades across networks are sufficient for the formation of echo chambers: An agent-based model

Jan-Philipp Fränken[1][*] **and Toby D. Pilditch**[2,3]

[1]Department of Psychology, University of Edinburgh, Edinburgh, EH8 9JZ, 7 George Square, United Kingdom
[2]School of Geography and the Environment, University of Oxford, OX1 3QY, South Parks Road, Oxford, United Kingdom
[3]Department of Experimental Psychology, University College London, WC1H 0AP, 26 Bedford Way, London, United Kingdom
[*]Correspondence and requests for materials should be addressed to jp.franken@ed.ac.uk

ABSTRACT

Investigating how echo chambers emerge in social networks is increasingly crucial, given their role in facilitating the retention of misinformation, inducing intolerance towards opposing views, and misleading public and political discourse (e.g., disbelief in climate change). Previously, the emergence of echo chambers has been attributed to psychological biases and inter-individual differences, requiring repeated interactions among network-users. In the present work we show that two core components of social networks —users self-select their networks, and information is shared laterally (i.e. peer-to-peer)— are causally sufficient to produce echo chambers. Crucially, we show that this requires neither special psychological explanation (e.g., bias or individual differences), nor repeated interactions — though these may be exacerbating factors. In fact, this effect is made increasingly worse the more generations of peer-to-peer transmissions it takes for information to permeate a network. This raises important questions for social network architects, if truly opposed to the increasing prevalence of deleterious societal trends that stem from echo chamber formation.

As we navigate social media platforms, we are free to customise our networks according to our individual needs: we choose to connect with users we like while we ignore others, we follow 'Influencers' that inspire us, and we selectively share and repost content. Combined with curated News Feeds, selective attention to and sharing of content has been associated with spreading of digital misinformation[1] and false news stories[2]. For example, it has been shown that people retweet false news stories more frequently than correct news stories, resulting in false rumour cascades on Twitter[2,3].

Over the past decade, several researchers investigated the spreading and retention of misinformation and false news on social media platforms[1,4–6] and their implications for e.g., the polarisation of opinions[7,8]. However, the scientific community still lacks clear answers to fundamental questions relating to 1) the general prevalence of misinformation and false news[9] and 2) their effects on individuals. Consequently, the demand for research investigating how misinformation and false news are spread through social media remains an important topic.

To understand the spreading of misinformation and false news, recent work has investigated the impact of echo chambers on digital misinformation. Echo chambers have been defined as enclosed epistemic systems where like-minded individuals reinforce their pre-existing beliefs[10]. The enclosing nature of echo chambers has been shown to induce intolerance towards opposing views[11] and quantitative analyses suggest that echo chambers may contribute to the spread of misinformation[1,12]. Corroborating these findings, recent empirical research has shown that echo chambers lead to a misleading of public and political discourse, such as disbelief in climate change[13,14]. Investigating how echo chambers emerge on social media thus offers an important opportunity for understanding and potentially counteracting the occurrence of digital misinformation.

Importantly, although some social media users might 'live' in echo chambers, using social media does not necessarily imply restricted exposure to information. Compared to non-users, the average social media user experiences more diverse content[15]. Moreover, recent theoretical work suggests that echo chambers might improve individual access to information via optimizing the allocation of information resources[16]. These findings further highlight the importance of clarifying how echo chambers emerge on social media, and raise questions about the value-free nature of echo-chambers. What factors lead to an echo-chamber causing harmful outcomes vs. beneficial ones?

In the present work, we formally study echo chamber formation within simulated populations of social media users. We expand the previous literature through two primary contributions, motivated by two important elements of social networks: 1) users are self-selecting in their peer networks (e.g., form connections based on friendship), and 2) users can share information rapidly in a peer-to-peer manner (i.e. lateral transmission). Specifically, we investigate the emergence of echo chambers as a

result of a single pass-through of information prior to repeated interaction between users (contribution 1). Moreover, we explore the impact of a social media user's perceived credibility of their peers on the formation of echo chambers (contribution 2).

We test the robustness of these contributions across a wide range of network setups varying in terms of the epistemic authority (expertise strength) of users (robustness check 1), the percentage of users sharing their beliefs with their peers (robustness check 2), and their connectivity density (robustness check 3). Moreover, we contrast two different hypothetical populations of agents (see Methods for details). For the first population ('social agents'), we assume that social media users select their network-peers based on whom they like (i.e. positive credibility). Conceptually, social agents are comparable to social media users that customise their networks based on positive perceptions of others (e.g., friendship). The second population ('asocial agents') functions as control check for the relevance of self-selection based on friendship / positive perceptions of others. Consequently, asocial agents select their network peers at random and independent of their perceived credibility. Given these considerations, we hope to further clarify the relationship between various conflated causes of echo chambers.

Background and motivation

Echo chambers

To investigate when and how echo chambers emerge, it is important to explore their causes. These might be routed in psychological biases: previous analyses of echo chambers and their impact on digital misinformation identified confirmation bias - seeking information confirming one's prior beliefs[17] - and social influence - peoples' tendency to align their behaviour with the demands of their social environment[18] - as key driving factors of echo chamber formation[1,8,19]. Similarly, work in statistical physics has shown that confirmation bias induces clustering of like-minded individuals (i.e. echo chambers) and proliferation of opinions[20].

The above findings might be explained by the fact that confirmation bias leads to selective avoidance of information challenging one's prior beliefs, and consequently, limited access to cross-cutting content on social media such as Facebook[6,20]. Along with psychological biases, it has also been argued that cognitive differences between individuals might induce echo chambers[21]. Overall, these findings suggest that both psychological variables and cognitive variability among individual agents might be necessary requirements for the formation of echo chambers.

Aiming to clarify the necessity of psychological variables and heterogeneity, recent simulation-based work has investigated echo chamber formation in an idealised population of homogeneous rational (i.e. Bayesian) agents engaging in repeated interaction[10,22]. Results provided a formal argument for the inherent susceptibility of social networks towards echo chamber formation despite *absence* of cognitive differences among agents. In other words, the findings of[10] suggest that the *structure* of social networks alone is sufficient for the formation of echo chambers.

Importantly, while previous work on echo chambers[10,22] and opinion dynamics (see also[23]) already shows that repeated interaction solidifies echo chambers, the present work investigates whether the initial way in which information arrives into the network already 'skews the pitch'. This is motivated by theoretical[24] and simulation-based[25] work on information cascades, which showed that single interactions can result in maladaptive collective outcomes prior to repeated interaction, and despite individually rational agents. Consequently, a single pass-through of information between generations of agents that update their beliefs sequentially might suffice for the emergence of echo chambers. In other words, the connective density of social networks—including lateral transmission of information and limited access to the knowledge of the entire network—may be sufficient for echo chamber formation, even before repeated interaction, and without recourse to cognitive differences / biased explanations.

To investigate this we employed an agent-based model (in line with prior research on related phenomena[10,22]) that allowed us to simulate whether echo chambers emerge in social agents as a consequence of social network structure (i.e. connective density - what percentage of the network is any given user *directly* connected to?) in the *absence* of repeated interaction (contribution 1). Specifically, we focused on a single cascade, meaning that all agents updated their beliefs sequentially through a single interaction. This method additionally allowed us to isolate structural aspects of a network from psychological variables and inter-individual differences among agents. Therefore, agents in our simulations were furnished with a homogeneous cognitive architecture, forming beliefs normatively through Bayesian updating.

Source credibility

The credibility of a source plays an important role when integrating their beliefs with our own observations and prior expectations[26,27]. Moreover, source credibility plays a critical role in persuasion and argumentation theory, especially in the context of politics[28–30], which has become increasingly influenced by online communication systems such as Facebook[31]. Both heuristic accounts, such as the heuristic-systematic model (HSM)[32] and dual-process theories, including the elaboration-likelihood model (ELM)[33] have been used to study the influence of credibility on persuasion, showing a positive general impact of credibility on persuasion[34] that has been extended to specific domains such as exercise intentions[35].

More recently, research has investigated the influence of source credibility from a Bayesian perspective, meaning that credibility is modeled as an analytic cue influencing the probability of accepting a message / updating a belief[36–39]. The Bayesian account provides a quantitative, normative framework for modeling belief updating under consideration of the influence of credibility, which is a key advantage over alternative models such as ELM and HSM where the credibility of sources is conceptualised as a heuristic cue, which does not allow for a quantitative formalisation. Empirical work supports the suitability of Bayesian representations of belief formation and credibility (see e.g.,[40]). Due to the quantitative nature of our analysis, we make use of the Bayesian credibility account, expanding previous simulation work[10,22] through a measure of the credibility of a source (i.e. an agent/social media user) in our model (contribution 2). The next section introduces the details of our agent-based model.

Model design and Simulations

Agent-based models

Agent-based models (ABMs) are simulated multi-agent systems that provide a formal framework for studying cognitive functions and behaviours in social environments —such as social media— which involve complex and dynamic interactions between users[41]. Of particular relevance for our work is the advantage that ABMs allow for the capturing of emergent phenomena on a network level which is not possible when studying individual users in isolation from their peers[42]. The advantages of ABMs for the study of complex and dynamic systems have been leveraged in several related studies, including ABMs of belief formation in the context of climate change[43] and micro-targeting[44]. A full discussion of ABMs and their advantages for the study of dynamic systems and emergent phenomena is beyond the scope of the present contribution. In the remainder of this paper, we will therefore only refer to aspects of ABMs that are relevant to our model. Further details about ABMs and their general advantages can be found elsewhere[42,45].

The three central components of ABMs are agents, patches, and links. Agents are the actors, and in our social network they correspond to individual users. Agents were furnished with cognitive functions and possible behaviours, including attention (detecting public declarations of others), belief revision (updating a prior belief-state based on observing another agent's belief), and declaration (commit to a belief using a decision rule). All agents were furnished with the same cognitive functions and possible behaviours. Links represent edges between agents. In the present model, (bidirectional) links were employed to enable signaling of public belief declarations between agents, and thus represent the (social) network connections. Patches are the building blocks of the environment in which agents act, though not relevant to the current model, like agents, patches can change dynamically. They have been used to model, for example, fluctuations in fish stock at a specific location[46].

Here, we reconstructed an idealised social network similar to[10,25] in which users (i.e. agents) form binary beliefs through a single interaction. As such, our agent-based model (ABM) captured agents' social environment and temporal dynamics of belief change, which both form necessary requirements for the observation of echo chambers. Additionally, our ABM accounted for heterogeneity between individual agents (such as different prior beliefs), which is an important general advantage of agent-based simulations[41]. Put another way, though we designed agents such that their cognitive *functions* are the same, we allowed for the diversity among agents of particular belief *values*.

Bayesian source credibility model

Bayesian theories of reasoning and decision making propose that a person's prior belief in a hypothesis is represented as subjective probability $P(h)$ taking values between 0 and 1 (see e.g.,[39,47]). Upon observing new data, d, the Bayesian framework posits that the posterior probability of a hypothesis, $P(h|d)$, is given by the normalized product of the likelihood $P(d|h)$ and the prior $P(h)$:

$$P(h|d) = \frac{P(d|h)P(h)}{\sum_{h'} P(d|h')P(h')}. \tag{1}$$

As in[10,22], we wanted to ensure that individual differences and psychological variables would not conflate with the impact of network structure on the formation of echo chambers. Consequently, all agents were idealised reasoners adhering to the principles of Bayesian updating. To include the credibility of a source, we used a Bayesian source credibility model (BSCM)[36–38]. In the BSCM framework, credibility has two components, perceived expertise, $P(e)$, and perceived trustworthiness, $P(t)$ (see also[40]). The perceived expertise of a source refers to the probability of the source's communicated belief being correct. The perceived trustworthiness of a communicating agent refers to the communicator's intention to tell the truth. In other words, $P(t)$ can be thought of as what the communicator believes to be true, independent of whether the belief is correct or not. $P(e)$ and $P(t)$ (which are orthogonal in BSCM) are both incorporated within the belief revision process of the present agent-based model (see Fig. 1).
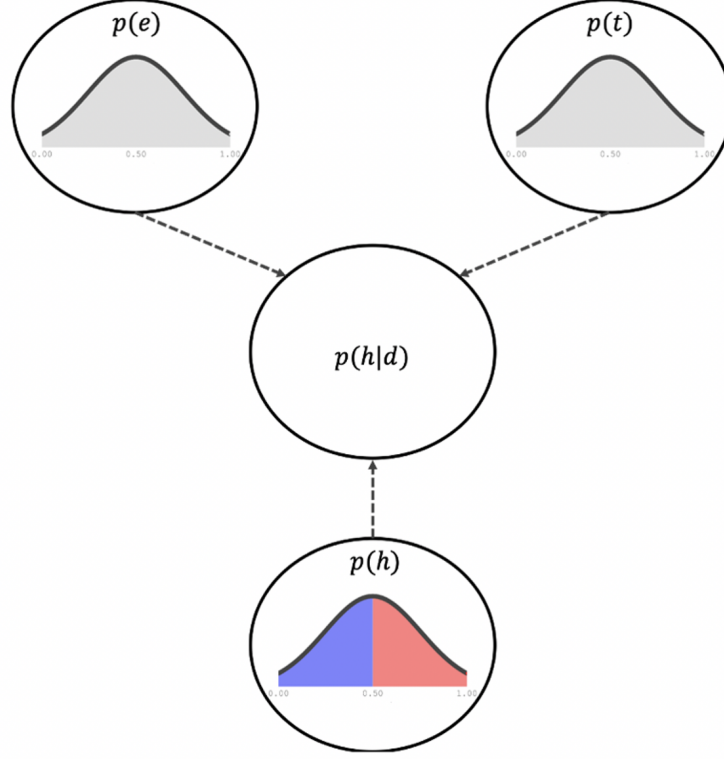
Figure 1. Figure shows how a prior belief and the perceived credibility of a source (operationalised via expertise and trustworthiness) interact within the BSCM framework. Values between 0.00 and 0.50 (blue) = $P(\neg h)$; Values between 0.51 and 1.00 (red) = $P(h)$.

In the BSCM framework, the likelihood for an agent that supports a hypothesis $P(h)$ is given by

$$P(d|h) = \sum_{e't'} P(h|e',t')P(e')P(t').$$ (2)

In the same way, the likelihood for an agent that does not support a hypothesis $P(\neg h)$ corresponds to

$$P(d|\neg h) = \sum_{e't'} P(\neg h|e',t')P(e')P(t')$$ (3)

(see[37,40], for further details).

The orthogonal nature of expertise and trust is typically operationalised such that trust being high or low leads to changes in the direction of belief revision (i.e. low trust makes you revise your beliefs in the opposite direction than high trust), whilst expertise moderates the strength (size) of the revision (see[36,40]). As a consequence, our manipulation of the perceived expertise strength, e, is a way to control how strongly a communication target is influenced by the expertise of a source (robustness check 1; see Methods). Conceptually, expertise strength can be compared of epistemic authority (i.e. knowledge superiority; see also[48]), which suggests that a source with higher expertise strength is going to exert stronger influence on a receiver during belief revision. To explore the impact of knowledge superiority, we investigated the influence of three different parameter settings for expertise strength on echo chamber formation (see Table 1).

Simulations

Agents were randomly assigned to x-y coordinates in a 2-dimensional space, forming links with their nearest neighbors (refer to[41], for further details on how to spatially allocate agents in ABMs). Proximity was measured in terms of Euclidean distance, which has been used as a proxy for relational proximity in social networks[49]. Prior to the start of a simulation, each agent

sampled their own prior belief $P(h)$ and subjective e and t values from univariate Gaussians with μ=0.5 and σ^2=0.20 (for all three estimates, distributions were truncated between [0, 1]). Thus, $P(h)$, e, and t differed heterogeneously within our agent population. Following revision of their prior belief according to BSCM, agents declared for one of the two beliefs based on a deterministic decision rule:

$$Belief = \begin{cases} h \text{ if } P(h|d) > 0.5 \\ \neg h \text{ if } P(h|d) < 0.5. \end{cases}$$

If $P(h) = 0.50$, an agent declared either belief with a probability of 50%. The declared belief (i.e. h or $\neg h$) was then made pubic based on the P(Declaration) probability (robustnes check 2) which was manipulated between simulations (see Table 1). For example, a declaration of 1 means all agents made their beliefs public, while 0.1 means that there is a 10% probability for each agent making their opinion public. The conceptual motivation for P(Declaration) is based on recent findings showing that *most* social media users do not discuss their political beliefs on social media, but mainly focus on exchanging shared hobbies and passions[50]. We wanted to ensure that our simulation results are robust across social networks varying in terms of the percentage of users discussing their beliefs with their peers. We thus explored the formation of echo chambers across three levels of P(Declaration) which can be thought of as a proxy for the percentage of people exchanging their beliefs about a particular topic (e.g., politics or news).

In addition to P(Declaration), we varied the connective density between simulations (i.e. what percentage of the overall network is any given user *directly* connected to) from 0.5% to 50% (robustness check 3). This allowed us to test whether the formation of echo chambers changes as a function of the number of "generations" it takes for information to pass across a network (i.e. how many round of peer-to-peer communication to get across the network). In this way, the more links (connected friends / users) an agent has, the more likely the agent is to see information appearing on the network early (and thus the shorter the cascade). This can also be compared to an increased number of friends on Facebook or followers on Twitter, where the information load of a user's News Feed increases if the person has more friends (Facebook) or follows more users (Twitter).

Name	Description	Levels
Connectivity density (%)	(Links per Agent / Total Number of Agents) * 100	0.5, 1.0, 1.5, ... 50.0
Expertise strength	Manipulate the magnitude of expertise influence	0.00, 0.10, 0.20
P(Declaration)	Probability of making a belief public	0.10, 0.50, 1.00

Table 1. Robustness checks.

To measure echo chambers effectively across simulations, we were first interested in measuring global proportions of beliefs across the whole network (i.e. the relative number of agents with belief h compared to agents entertaining belief $\neg h$). Based on previous simulation-based work, we expected that global proportions would consistently approximate 50/50 across both agent populations[25]. This measure was necessary to ensure that echo chambers are not a by-product of a dominant network-wide belief. Following checks for possible network-wide belief confounds, our key dependent measure of echo chamber formation was the average percentage of like-minded neighbours an agent possessed (i.e. the *local* network similarity). This measure corresponded to the average percentage of agents in the target's direct network that shared the same belief as the target. For example, 50% means that, on average, agents had equal proportions for each belief type in their direct network, where direct network refers to the fraction of the whole network that is directly connected to an agent.

In line with previous work[25], we expected that social agents would show increased percentages of like-minded neighbours (i.e. echo chambers) for low connectivity density values. We did not expect clustering effects in the asocial population in which agents selected network-peers at random and computed stochastic credibility estimates for a communicating source (see Methods for details). Each system specification (connectivity density (100) x expertise strength (3) x P(Declaration) (3)) was ran independently 50 times, taking an average set of values for each specification. The total number of agents (n = 1000) was consistent across simulations. Simulations were conducted independently for each of the two agent populations.

Results

Fig. 2 summarizes our central findings. We predicted that for both social and asocial agents, global belief proportions would consistently approximate 50/50. Fig. 2A (social agent population) and Fig. 2B (asocial agent population) confirm these predictions. Specifically, Fig. 2A and 2B show that global proportions of beliefs in both populations consistently approximated 50/50 irrespective of varying P(Declaration), connectivity density, or expertise strength. This finding is important, as it ensures that potential clusters of like-minded others (Fig. 2C-D) do not result from a global bias towards either belief.

For social agents, the average proportion of like-minded neighbors (Fig. 2C) increased as a function of increasing expertise strength, increasing P(Declaration) and decreasing connectivity density (x-axis). Importantly, to fully reduce the formation of echo chambers, the average network member must be connected to around 15-20% of the network, which is infeasible considering the size of real world social networks which can have several Billion users. Our results are comparable to previous simulation-based work exploring echo chamber formation in a a population of stochastic reinforcement learners[25], although the present findings revealed less intense clustering effects.
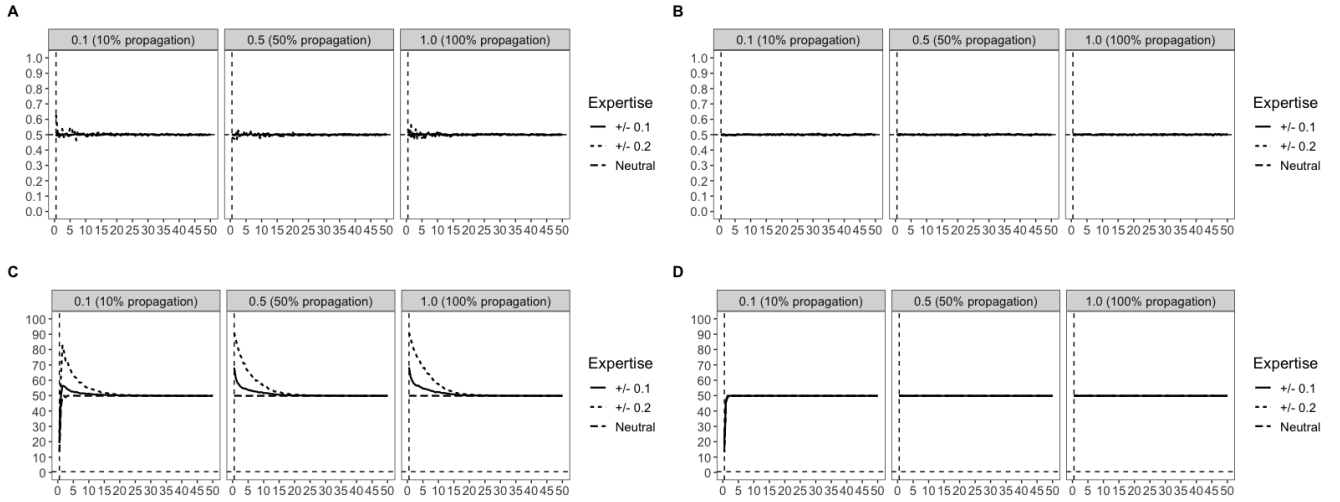

Figure 2. Main results. A: global belief proportions, social agents. B: global belief proportions, asocial agents. C: average percentage of like-minded neighbors (i.e. echo chambers), social agents. D: average percentage of like-minded neighbors (i.e. echo chambers), asocial agents.

The finding that an expertise strength of 0 (i.e. neutral) prevented the formation of echo chambers is a natural result of our model. Specifically, setting expertise strength to 0 reduces the communicative impact of a source to 0, irrespective of their perceived credibility (for details see Methods). Consequently, a receiver won't be influenced by the communicating source, which can (conceptually) be compared to disregarding the belief of a social media user that has no epistemic authority (i.e., no knowledge about the topic of discussion).

Fig. 2D shows the clustering results from our asocial agent population. If agents randomly connect with their network-peers, meaning that they stochastically sampled credibility estimates for each source, no echo chambers emerged. This finding highlights the importance selecting network-peers based on positive credibility perceptions (e.g., friendship): while stochastic selection prevents echo chambers, selecting network peers based on whom one likes is a key requirement for echo chamber formation.

The results in the left panels of Figs. 2C-D (P(Declaration) = 10%) showing a reduced clustering effect for connectivity density values of 0.5% can be fully attributed to fracturing of the network (i.e. no information cascade occurred). This finding demonstrates that if we did not share information on social media, we would not get information cascades. To visualise the above findings, Fig. 3 includes example outcomes of post-cascade belief proportions with 1% and 5% connectivity density. A and C correspond to our asocial agent population and B and D pertain to the asocial population.

Discussion

The aim of this work was to disentangle previously conflated causes of echo chambers on social media. Specifically, we examined whether echo chambers emerge in a population of homogeneous, individually rational users that engage in a single interaction. Our results show that echo chambers emerge in a population of individually rational agents who integrate the beliefs of others with their own prior beliefs in a Bayesian manner. Moreover, we showed that echo chambers emerge through a single pass-through of information prior to repeated interaction.

These results suggest that previously identified causes of echo chambers, including psychological biases and inter-individual differences in terms of the cognitive architecture of agents, might not be necessary for the observation of echo chambers. Moreover, our findings suggest that repeated interaction, which has been a key component in previous models of echo chambers[10,22], might not be required for the observation of echo chambers. On average, each belief was equally represented in
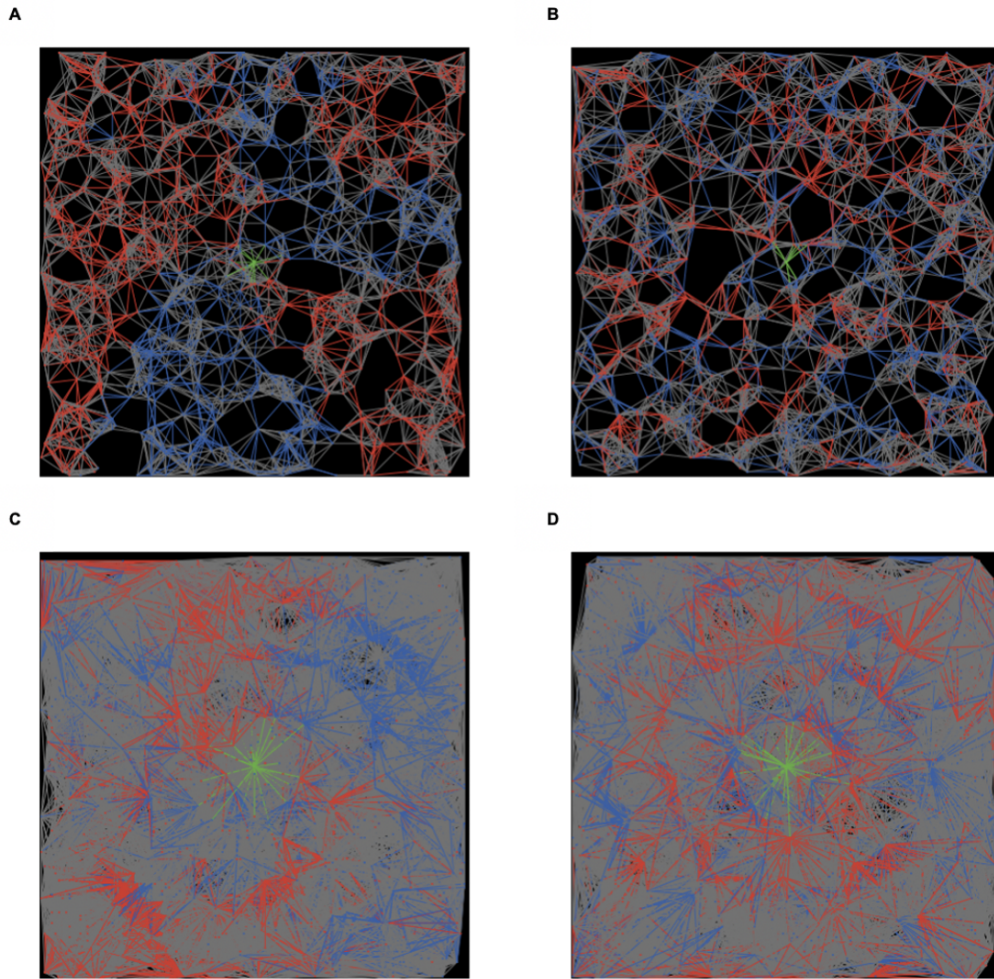
Figure 3. Example post-cascade networks. Grey = agents that did not declare a belief; blue = agents believing $\neg h$; red = agents believing h. Connectivity density values of upper networks: A (social agents): 1%; B (asocial agents): 1%. Connectivity density values of lower networks: C (social agents): 5%; D (asocial agents): 5%.

our simulations. Thus, our results further show that segregated groups did not evolve as a consequence of a global dominance of a particular belief (see Figs. 2-3).

Furthermore, and central to the present paper is the finding that the emergence of echo chambers was confined to the social agent population in which network-peers were selected on the basis of positive credibility estimates. Specifically, we showed that self-selecting peers based on whom one likes (i.e. perceives positively) carries a potentially deleterious consequence for the formation of echo chambers. If network-peers are selected at random and thus independent of whom one likes or perceives positively (asocial agents), no echo chambers emerged.

Overall, our findings illustrate that echo chambers, which might induce spreading and retention of misinformation[2], conspiratorial thinking[1], and political polarisation[7,51], are not necessarily caused by the inhabitants of social networks directly. Rather, the *structure* of social networks, and notably the constraining lateral (i.e. peer-to-peer) transmission of information can be sufficient for echo chamber formation. The degree of making opinions public (P(Declaration)) did affect echo chamber formation only if it was so low that it effectively fractured the functional message passing around the network. Additionally, the magnitude of expertise strength modulated the influence of credibility, resulting in increased echo chamber effects for higher levels of expertise. This suggests that being friends with expert users might exacerbate the formation of echo chambers. Given that the present simulations included rational Bayesian agents, it is further expected that incorporation of additional psychological variables, such as the confirmation bias[1,19,20], might intensify the strength and persistence of echo chambers (see also[25]).

More generally, our results show that agent-based models, which enable capturing of dynamic interactions between

individuals, provide a valuable opportunity for studying the formation of emergent phenomena such as echo chambers. This is in line with a growing body of literature employing agent-based models to investigate several related phenomena, including opinion polarisation[49], identity search[52], disbelief in climate change[43] and micro-targeting[44]. Given the potential of agent-based models for studying echo chambers and related emergent phenomena, it is important that further work develops interventions which might reduce the occurrence of opinion segregation. Such interventions might extend previous work suggesting that 'educational broadcasts' across a network of idealised Bayesian agents (similar to the present model) reduce echo chamber formation[10].

An important avenue for further work might be a closer examination of the belief updating process of the present agent populations. Specifically, agents in the present populations updated beliefs sequentially based on the declarations of previous generations. Recent empirical studies demonstrated that people are sensitive to such statistical dependencies in social learning. Across both abstract (e.g., estimating the number marbles in an urn; see[53]) and more applied tasks (e.g., rating the suitability of a political candidate for public office), it has been shown that people differentiate between the evidential signal of independent beliefs and dependent beliefs that were formed sequentially. If beliefs of others were formed sequentially, people updated their prior beliefs less. Considering these findings, an important step for follow-up simulations involves testing the robustness of echo chambers between varying levels of belief dependencies in a network. Here, empirical findings by[53] might serve as input validation for simulations, enabling agents to differentiate between independent and sequentially updated beliefs.

In summary, while the study of echo chambers does by no means exhaustively addresses the issue of digital misinformation, it provides an important contribution to our understanding of how people settle on beliefs in social networks. Here, we show that echo chambers emerge as a consequence of self-selection of network peers (e.g., friends), limited access to information (i.e. local connectivity), and lateral transmission of information between social agents. This suggests that social networks by design can be causally sufficient for echo chamber formation. As a consequence, we show that previously conflated causes such as psychological biases, inter-individual differences, and repeated interaction, might not be necessary. Rather, such inter-individual differences and psychological variables, including confirmation bias, are expected to exacerbate the formation of echo chambers. Finally, we show that the computations involved in evaluating the credibility of a source play a key role in the formation of echo chambers. Together, these findings raise questions for social network architects, and notably the potentially deleterious consequences of encouraging peer-to-peer content sharing.

Methods

BSCM details

Table 2 shows the conditional probability table which specified how the different components of the likelihood were computed. To ensure that the direction of the influence of expertise strength (**e**) matched an agent's prior belief (i.e. towards 1 if $P(h) > 0.5$ and towards 0 if $P(h) < 0.5$), we flipped the impact of expertise strength based on the prior. Similar to an indicator function, I thus returned 1 for agents having a prior belief $P(h) > 0.5$ and -1 for agents having a prior belief $P(h) < 0.5$. τ is an additional constant that quantifies the presence vs. absence of expertise. Manipulating values of τ results in stronger differences between varying levels of expertise strength. For all simulations, $\tau = 2$, meaning that the magnitude of belief change for a medium expert (i.e. **e** = 0.1) ranged from -0.2 to 0.2 while the magnitude for a high expert (i.e. **e** = 0.2) ranged from -0.4 to 0.4. For **e** = 0, there was no effect on belief change.

	e,t	¬e,t	e,¬t	¬e,¬t
h	$0.5 + \text{eI}_{P(h)>0.5} * \tau$	$0.5 + \text{eI}_{P(h)>0.5}$	$0.5 - \text{eI}_{P(h)>0.5} * \tau$	$0.5 - \text{eI}_{P(h)>0.5}$
¬h	$1 - (0.5 + \text{eI}_{P(h)>0.5} * \tau)$	$1 - (0.5 + \text{eI}_{P(h)>0.5})$	$1 - (0.5 - \text{eI}_{P(h)>0.5} * \tau)$	$1 - (0.5 - \text{eI}_{P(h)>0.5})$

Table 2. Conditional probability table.

The BSCM architecture provides a general framework for updating one's belief under consideration of a source's credibility. It does not specify how individual agents estimate the perceived expertise and perceived trustworthiness values of a source during interaction. Here, we refer to an important element of social networks: selecting network-peers on the basis of whom one likes (i.e. positive credibility estimates). We operationalised this self-selection property in a hypothetical population of 'social agents' who weigh the perceived credibility of a communicating source based on the beliefs of their peers. Specifically, following observation of a source's belief in a hypothesis, a target agent (receiver) weighs the expertise and trustworthiness of a communicating source based on the proportion of agents in the receiver's direct network entertaining the same belief as the source and the number of agents in the network entertaining the opposite belief of the source. More formally, this can be

written as

$$P(e) = \frac{\sum_{i=1}^{N_h} e_i}{\sum_{i=1}^{N_h} e_i + \sum_{i=1}^{N_{\neg h}} e_i} \tag{4}$$

$$P(t) = \frac{\sum_{i=1}^{N_h} t_i}{\sum_{i=1}^{N_h} t_i + \sum_{i=1}^{N_{\neg h}} t_i} \tag{5}$$

where e and t correspond to the subjective estimates for individual agents in the receiver's network (sampled from univariate Gaussians, see Fig. 1).

We contrasted the above population of social agents with a second population of asocial agents in which connections with peers were established at random. Random selection of peers was operationalised by random sampling of perceived expertise and perceived trustworthiness of sources from a uniform distribution bounded between [0, 1]. This means that asocial agents computed stochastic credibility estimates for a source irrespective of the beliefs of their network peers. Asocial agents thus enabled to assess the relevance of selecting network-peers based on positive credibility estimates of others / friendship. For simulations across both hypothetical populations (social and asocial), perceived expertise $P(e)$ and trustworthiness $P(t)$ estimates were plugged into Equations 2-3 to compute the likelihood of observing a source's belief given the receivers prior belief and the credibility of the source.

More on simulations

The model was implemented in NetLogo version 6.0.4[54]. All simulations were performed in R using the package `RNetLogo()`[55]. In the following, we outline the basic technical details of our model. Algorithm 1 shows the setup procedure of the network, including placing agents in space, sampling prior $P(h)$, e, and t values, and forming connections with peers based on proximity as measured by Euclidean distance (a proxy for relational proximity in social networks; see[49]).

Algorithm 1 Setup network

1: **procedure** PLACE AGENTS
2:     Create N agents
3:     **for** $i = 1$ to N **do**
4:         set i position random x-y coordinate
5:     **end for**
6: **end procedure**
7: **procedure** SETUP PRIORS AND LINKS
8:     **for** $i = 1$ to N **do**
9:         $P(h)[i] \leftarrow x \sim N(\mu, \sigma^2)$
10:        $e[i] \leftarrow x \sim N(\mu, \sigma^2)$
11:        $t[i] \leftarrow x \sim N(\mu, \sigma^2)$
12:        create links with n nearest neighbors $\triangleright$ based on Euclidean distance
13:     **end for**
14: **end procedure**

Simulations were initiated through a neutral event node. Neutral in this case means that the first agent sampled randomly from the set of beliefs $\{h, \neg h\}$ each time it communicated to a target. Due to this stochastic process, on average, half of the 1st generation agents receiving input from the neutral event node should arrive at belief h while the other half will arrive at belief $\neg h$ after BSCM integration. Specifically, this included attending to the declared beliefs of the neutral event node, revising initial beliefs based on the perceived credibility of the neutral event node, and deciding whether to declare an opinion public according to P(Declaration).

The number of agents the neutral event node communicated to was determined by the connectivity density. Connectivity values above 50% were omitted as this would have enabled every other agent to be connected to the neutral event node in the 1st generation, precluding the occurrence of a cascade. Similarly, values below 0.5% would have resulted in fracturing of the network. To improve the readability of plotted example networks (Fig. 3), the initial neutral event node was placed in the center of each simulation (i.e. central x-y-coordinate).

After revising their prior beliefs, the first-generation agents (those that received input from the neutral event node) made their beliefs public based on the manipulated P(Declaration) value. Their communication targets (i.e. second-generation) then used the communicated opinion of the first-generation agents as input for their own belief revision following the same procedure. Algorithm 2 shows the basic steps involved in a single instance of belief revision (i.e. from one generation to the the next). Here, source refers to an agent from the previous generation that already publicly declared a belief (i.e. h or $\neg h$) and the connections of the source refer to the potential communication targets of the next generation. As we investigated whether a single pass-through of information (i.e. single interaction) was sufficient for the formation of echo chambers, we did not allow for repeated interaction, meaning that agents could not qualify as communication targets after declaring for a belief.

Algorithm 2 Updating beliefs

1: **procedure** SOURCE SELECTS COMMUNICATION TARGET
2:   **if** matchCounter $\leq$ 1 **then**
3:     **for** $i = 1$ to n_{source} **do** $\triangleright$ n = source's connections
4:       **if** i = neutral **then** $\triangleright$ check if target did not already declare a belief
5:         belief[i] = BSCM(source, n[i]) $\triangleright$ n = target's connections
6:         **if** random(0.01,1.00) > P(Declaration) **then**
7:           propagate belief to next generation
8:         **end if**
9:       **end if**
10:     **end for**
11:   **end if**
12: **end procedure**

The process of transmitting beliefs continued down successive generations until the network was either completely saturated (i.e. all agents committed to a belief) or the number of believers (i.e. h or $\neg h$) did not change for two consecutive time periods. Algorithm 3 illustrates how this procedure was implemented in our model.

Algorithm 3 Stop simulation

1: **procedure** CHECK IF NETWORK IS SATURATED OR FRACTURED
2:   matchCounter = 0 $\triangleright$ counts how often no change occurred between generations
3:   $Count_h = 0$
4:   $Count_{\neg h} = 0$
5:   **for** $i = 1$ to N **do** $\triangleright$ N = all agents
6:     **if** belief[i] = h **then**
7:       $Count_h = Count_h + 1$
8:     **end if**
9:     **if** belief[i] = $\neg h$ **then**
10:       $Count_{\neg h} = Count + 1$
11:     **end if**
12:   **end for**
13:   **if** $Count_h = Count_{\neg h}$ **then**
14:     matchCounter = matchCounter + 1
15:   **end if**
16: **end procedure**

References

1. Del Vicario, M. *et al.* The spreading of misinformation online. *Proc. Natl. Acad. Sci.* **113**, 554–559 (2016).

2. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151 (2018).

3. Dizikes, P. Study: On twitter, false news travels faster than true stories. *Pobrane z: http://news. mit. edu/2018/study-twitter-false-news-travels-faster-true-stories-0308 (20.04. 2018)* (2018).

4. Starbird, K., Maddock, J., Orand, M., Achterman, P. & Mason, R. M. Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing. *IConference 2014 Proc.* (2014).

5. Bessi, A. *et al.* Science vs conspiracy: Collective narratives in the age of misinformation. *PloS one* **10**, e0118093 (2015).

6. Bakshy, E., Messing, S. & Adamic, L. A. Exposure to ideologically diverse news and opinion on facebook. *Science* **348**, 1130–1132 (2015).

7. Bessi, A. *et al.* Users polarization on facebook and youtube. *PloS one* **11** (2016).

8. Sikder, O., Smith, R. E., Vivo, P. & Livan, G. A minimalistic model of bias, polarization and misinformation in social networks. *Sci. Reports* **10**, 1–11 (2020).

9. Lazer, D. M. *et al.* The science of fake news. *Science* **359**, 1094–1096 (2018).

10. Madsen, J. K., Bailey, R. M. & Pilditch, T. D. Large networks of rational agents form persistent echo chambers. *Sci. reports* **8**, 12391 (2018).

11. Takikawa, H. & Nagayoshi, K. Political polarization in social media: Analysis of the "twitter political field" in japan. In *2017 IEEE International Conference on Big Data (Big Data)*, 3143–3150 (IEEE, 2017).

12. Törnberg, P. Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLoS one* **13** (2018).

13. Jasny, L., Waggle, J. & Fisher, D. R. An empirical examination of echo chambers in us climate policy networks. *Nat. Clim. Chang.* **5**, 782–786 (2015).

14. Jasny, L. & Fisher, D. R. Echo chambers in climate science. *Environ. Res. Commun.* **1**, 101003 (2019).

15. Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. & Kleis Nielsen, R. Reuters digital news report 2017. *Reuters Inst. for Study Journalism, Univ. Oxford. Link: https://reutersinstitute. politics. ox. ac. uk/sites/default/file %%20Report* **202017** (2017).

16. Jann, O. & Schottmüller, C. Why echo chambers are useful. Tech. Rep. (2018).

17. Nickerson, R. S. Confirmation bias: A ubiquitous phenomenon in many guises. *Rev. general psychology* **2**, 175–220 (1998).

18. Kelman, H. C. Compliance, identification, and internalization three processes of attitude change. *J. conflict resolution* **2**, 51–60 (1958).

19. Starnini, M., Frasca, M. & Baronchelli, A. Emergence of metapopulations and echo chambers in mobile agents. *Sci. reports* **6**, 31834 (2016).

20. Ngampruetikorn, V. & Stephens, G. J. Bias, belief, and consensus: Collective opinion formation on fluctuating networks. *Phys. Rev. E* **94**, 052312 (2016).

21. Barkun, M. *A culture of conspiracy: Apocalyptic visions in contemporary America*, vol. 15 (Univ of California Press, 2013).

22. Madsen, J. K., Bailey, R. & Pilditch, T. D. Growing a bayesian conspiracy theorist: An agent-based model. *Proc. Annu. Meet. Cogn. Sci. Soc.* (2017).

23. Lorenz, J. Continuous opinion dynamics under bounded confidence: A survey. *Int. J. Mod. Phys. C* **18**, 1819–1838 (2007).

24. Bikhchandani, S., Hirshleifer, D. & Welch, I. A theory of fads, fashion, custom, and cultural change as informational cascades. *J. political Econ.* **100**, 992–1026 (1992).

25. Pilditch, T. D. Opinion cascades and echo-chambers in online networks: A proof of concept agent-based model. *Proc. Annu. Meet. Cogn. Sci. Soc.* (2017).

26. Cuddy, A. J., Glick, P. & Beninger, A. The dynamics of warmth and competence judgments, and their outcomes in organizations. *Res. organizational behavior* **31**, 73–98 (2011).

27. Fiske, S. T., Cuddy, A. J. & Glick, P. Universal dimensions of social cognition: Warmth and competence. *Trends cognitive sciences* **11**, 77–83 (2007).

28. Housholder, E. E. & LaMarre, H. L. Facebook politics: Toward a process model for achieving political source credibility through social media. *J. Inf. Technol. & Polit.* **11**, 368–382 (2014).

29. Robinson, J. P., Shaver, P. R. & Wrightsman, L. S. *Measures of political attitudes.* (Academic Press, 1999).

30. Cialdini, R. B. & Cialdini, R. B. *Influence: The psychology of persuasion* (Morrow New York, 1993).

31. Bail, C. A. Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media. *Proc. Natl. Acad. Sci.* **113**, 11823–11828 (2016).

32. Chaiken, S. The heuristic—systematic. *Dual-process theories social psychology* 73 (1999).

33. Petty, R. E. & Cacioppo, J. T. The elaboration likelihood model of persuasion. In *Communication and persuasion*, 1–24 (Springer, 1986).

34. Chaiken, S. & Maheswaran, D. Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgment. *J. personality social psychology* **66**, 460 (1994).

35. Jones, L. W., Sinclair, R. C. & Courneya, K. S. The effects of source credibility and message framing on exercise intentions, behaviors, and attitudes: An integration of the elaboration likelihood model and prospect theory 1. *J. applied social psychology* **33**, 179–196 (2003).

36. Bovens, L., Hartmann, S. *et al. Bayesian epistemology* (Oxford University Press on Demand, 2003).

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

38. Harris, A. J. & Hahn, U. Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**, 1366 (2009).

39. Oaksford, M., Chater, N. *et al. Bayesian rationality: The probabilistic approach to human reasoning* (Oxford University Press, 2007).

40. Harris, A. J., Hahn, U., Madsen, J. K. & Hsu, A. S. The appeal to expert opinion: quantitative support for a bayesian network approach. *Cogn. Sci.* **40**, 1496–1533 (2016).

41. Wilensky, U. & Rand, W. *Introduction to agent-based modeling: modeling natural, social and engineered complex systems* (The MIT Press, Cambridge MA, 2015).

42. Madsen, J. K., Bailey, R., Carrella, E. & Koralus, P. Analytic versus computational cognitive models: Agent-based modeling as a tool in cognitive sciences. *Curr. Dir. Psychol. Sci.* **28**, 299–305 (2019).

43. Lewandowsky, S., Pilditch, T. D., Madsen, J. K., Oreskes, N. & Risbey, J. S. Influence and seepage: An evidence-resistant minority can affect public opinion and scientific belief formation. *Cognition* **188**, 124–139 (2019).

44. Madsen, J. K. & Pilditch, T. D. A method for evaluating cognitively informed micro-targeted campaign strategies: An agent-based model proof of principle. *PLoS one* **13**, e0193909 (2018).

45. Macal, C. M. & North, M. J. Tutorial on agent-based modeling and simulation. In *Proceedings of the Winter Simulation Conference, 2005.*, 14–pp (IEEE, 2005).

46. Bailey, R. M. *et al.* A computational approach to managing coupled human–environmental systems: the poseidon model of ocean fisheries. *Sustain. Sci.* **14**, 259–275 (2018).

47. Hahn, U. & Oaksford, M. The rationality of informal argumentation: a bayesian approach to reasoning fallacies. *Psychol. review* **114**, 704 (2007).

48. Walton, D. *Appeal to expert opinion: Arguments from authority* (Penn State Press, 2010).

49. Duggins, P. A psychologically-motivated model of opinion change with applications to american politics. *J. Artif. Soc. Soc. Simul.* **20**, 13 (2017).

50. Newman, N., Fletcher, R., Kalogeropoulos, A. & Nielsen, R. *Reuters institute digital news report 2019*, vol. 2019 (Reuters Institute for the Study of Journalism, 2019).

51. Del Vicario, M. *et al.* Echo chambers: Emotional contagion and group polarization on facebook. *Sci. reports* **6**, 37825 (2016).

52. Watts, D. J., Dodds, P. S. & Newman, M. E. Identity and search in social networks. *science* **296**, 1302–1305 (2002).

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

54. Wilensky, U. *et al.* Netlogo (1999).

55. Thiele, J. C. R marries netlogo: introduction to the rnetlogo package. *J. Stat. Softw.* **58**, 1–41 (2014).

Acknowledgements

Author contributions statement

J.-P.F. and T.D.P. designed research; J.-P.F. and T.D.P developed the agent-based model and conducted simulations; J.-P.F. and T.D.P. wrote the paper.

Additional information

Competing interests: The authors declare no competing interests.
Data availability: The model and data sets generated and analysed during the current study will be available in the accompanying GitHub repository.