**Title:**   
Interdependent Diffusion: The social contagion of interacting beliefs

**Author:**   
James P. Houghton

**Affiliation:**

Sloan School of Management

Massachusetts Institute of Technology

100 Main Street

Cambridge, MA 02142

**Please direct correspondence to**:

James P. Houghton,

19 Kimball Street, Cambridge MA 02140, USA

+1 281 728 6848

[houghton@mit.edu](mailto:houghton@mit.edu)

<https://orcid.org/0000-0002-6907-6973>

The author declares no competing interests.

**Keywords:**Social Contagion, Belief Systems, Social Learning, Social Networks, Collective Cognition, Polarization, Interdependent Diffusion

**Abstract:**

It is well known that humans prefer to adopt beliefs or ideas that are consistent with those they already hold. However, studies of social contagion generally ignore interactions between beliefs and assume that beliefs spread independently of one another. Is this a harmless simplification? Or does omitting interdependence between diffusants suppress important dynamics, and change the outcome of social contagion? Using carefully matched simulations of independent and interdependent diffusion, I identify two mechanisms unique to interdependent diffusion that allow worldviews to emerge without reference to any external truth, and allow polarization to develop in otherwise homogenous populations. These predictions were confirmed in a controlled laboratory experiment with 2400 participants in 120 artificial social networks. I conclude that the assumption of independence between diffusants is not as universally appropriate as its ubiquity would suggest. Instead, interdependence between diffusants is likely to be both common and consequential.

**Introduction:**

On June 18th, 2015, the U.S. Treasury announced that a portrait of a woman would appear on the ten-dollar bill. The same day, news emerged of a mass shooting at a historically black church in Charleston, South Carolina. Within twenty-four hours both stories had spread through the U.S. population, driven by broadcast news and social media. In working to understand the spread of news about the shooting, researchers would be justified in ignoring the news about the $10 bill, and vice versa. We can assume independence between these diffusants because the probability that an individual will choose to share one piece of news is not likely to be causally influenced by whether they have absorbed and shared the other.

In the days that followed the shooting, two other news items emerged on social media. The first was a report that the Charleston shooter had been motivated by racial hatred, which he found symbolized in the Confederate flag. The second was a call to remove the Confederate flag from the South Carolina state capitol grounds *(1)*. Even though these are distinct ideas, each spreading through a process of social contagion, we cannot ignore one in trying to understand the diffusion of the other. If an individual has previously adopted the belief that the flag should be removed from the capitol grounds, they will be more likely to believe that the shooter’s identification with the flag is politically relevant, and vice versa. Rather than being independent diffusants, these beliefs are *inter*dependent.

With good reason, the overwhelming majority of social contagion research over the last 50 years has assumed that diffusants spread independently of one another. Independence is an extremely useful and generative simplification. By assuming that diffusants do not interact, researchers can study the effects of social network structure, homophily, social reinforcement, or demographics on each contagion process in isolation. Researchers then linearly superimpose any additional contagion processes upon the first.

The independence assumption makes it easier to develop and communicate theories of social contagion. To highlight a few high-profile examples, when Granovetter showed that one’s distant acquaintances may provide more novel information than one’s close friends, he ignored how the receiver acquires the contextual knowledge needed to understand foreign information *(2)*. Similarly, when he showed how individuals’ heterogeneous responses to peer pressure can shape cascades of collective action, he set aside the social contagion of tactics, risk assessments, and political attitudes that accompany protest *(3)*. When Schelling showed that even weak preferences for similar neighbors could lead to segregation, he measured “similarity” along a single dimension *(4)*. When Watts highlighted the importance of local network structure for the initial germination of adoption cascades, he limited his model to one diffusant at a time *(5)*. When Burt showed that individuals who bridge a gap between strangers can benefit from brokering information between them, he assumed that all parties share enough common understanding to make the information transferable in the first place *(6)*. When Kempe, Kleinberg, and Tardos worked out how to promote a belief to maximize its spread, they ignored the simultaneous spread of other beliefs in the network *(7)*. As a final example, when Centola and Macy developed the theory of complex contagion, they specified that social reinforcement of the *same* behavior was required for individuals to adopt a behavior, omitting the possibility that related behaviors may also support adoption *(8)*.

Almost all empirical studies of social contagion have considered diffusants to be independent of one another. For example, Travers and Milgram’s use of chain letters to study social network structure may have reached different conclusions if individuals’ willingness to forward the letter had been influenced by other pieces of mail *(9)*. Lorenz *et al.* *(10)* and Muchnik, Aral, and Taylor *(11)* showed that social influence undermined the “wisdom of crowds” effect when the crowd assessed one item at a time. Rand, Arbesman, and Christakis *(12)* and Suri and Watts *(13)* showed that social network structure influenced the spread of cooperation, a single social behavior*.*

The independence assumption makes for parsimonious theory, and it reduces the complexity and expense of experiments. However, scholars of social contagion assume independence between diffusants so frequently and to such productive ends that they do not always make this assumption explicit. This is dangerous, in that readers – and occasionally authors themselves – may not recognize that an assumption is being made.

“Interdependent diffusion” describes any social contagion process in which individuals’ likelihood of adopting diffusant A is a function of their current state of adoption of B (C, D, etc.) *and* in which their likelihood of adopting B (C, D, etc.) is a function of their state of adoption of A. In the social contagion literature, only a few studies explicitly allow for this type of interaction between diffusants. Baldassari and Bearman *(14)* and DellaPosta, Shi, and Macy *(15)* both explore interdependence that arises when individuals’ opinions and preferences serve to signal membership in a particular group. In these models, various preferences tend to become popular in different parts of the social network, and positions on political issues become associated with lifestyle choices. Friedkin et al. *(16)* and Goldberg and Stein *(17)* explore interdependence that arises from logical interactions between diffusants and simulate the emergence of polarization or consensus. While they do not explicitly compare independent and interdependent diffusion, and they have not tested their theories empirically, these papers[[1]](#footnote-2) suggest that scholars of social contagion should reassess the common assumption of independence between diffusants.

This paper begins such a reassessment. How much does interdependence *matter*? In addressing this question, I argue that interdependence matters to scholars and practitioners if it 1) generates new sociological processes, 2) suggests new observable outcomes, and 3) has practical consequences for communication and social policy. If these results are not observed then the assumption of independence is sufficient. If they are observed, much work remains to understand how interdependence changes our understanding of social contagion.

This paper uncovers two new sociological processes that are unique to interdependent diffusion and which cannot be reduced to the familiar influences of network structure, homophily, social reinforcement, or demographics. First, when beliefs[[2]](#footnote-3) support one another’s adoption, they can “snowball” through a population to reach a far broader audience than any may have reached on its own. Secondly, when individuals have similar belief sets, they are more likely to respond in the same way to new beliefs to which they are exposed, and so become yet more similar.

Simulations in this paper predict that shared “worldviews” will emerge spontaneously from the process of interdependent diffusion. Specifically, subsets of a population will come to share a set of interconnected beliefs (and reject others that are equally available) without reference to any ground truth. Interdependent diffusion is also predicted to foment polarization by increasing similarity within ideological camps and difference between camps, and aligning the population along a “left-right” political axis.

Finally, this paper reports an experimental test of the above predictions that involved 2400 participants in 120 artificial social networks. Each individual was given clues to a mystery and was asked to share any promising leads with their social network neighbors. After eight minutes, they were asked to judge who performed the crime and how. The presented information was manipulated to create a world with strong interdependence between clues, and a comparison world in which the same clues were independent of each other. Unbeknownst to the participants, the mystery had no solution and the clues were perfectly symmetric and pretested to eliminate outside bias. Regardless, participants confidently surmised an answer, and in so doing confirmed the predicted emergence of worldviews and polarization.

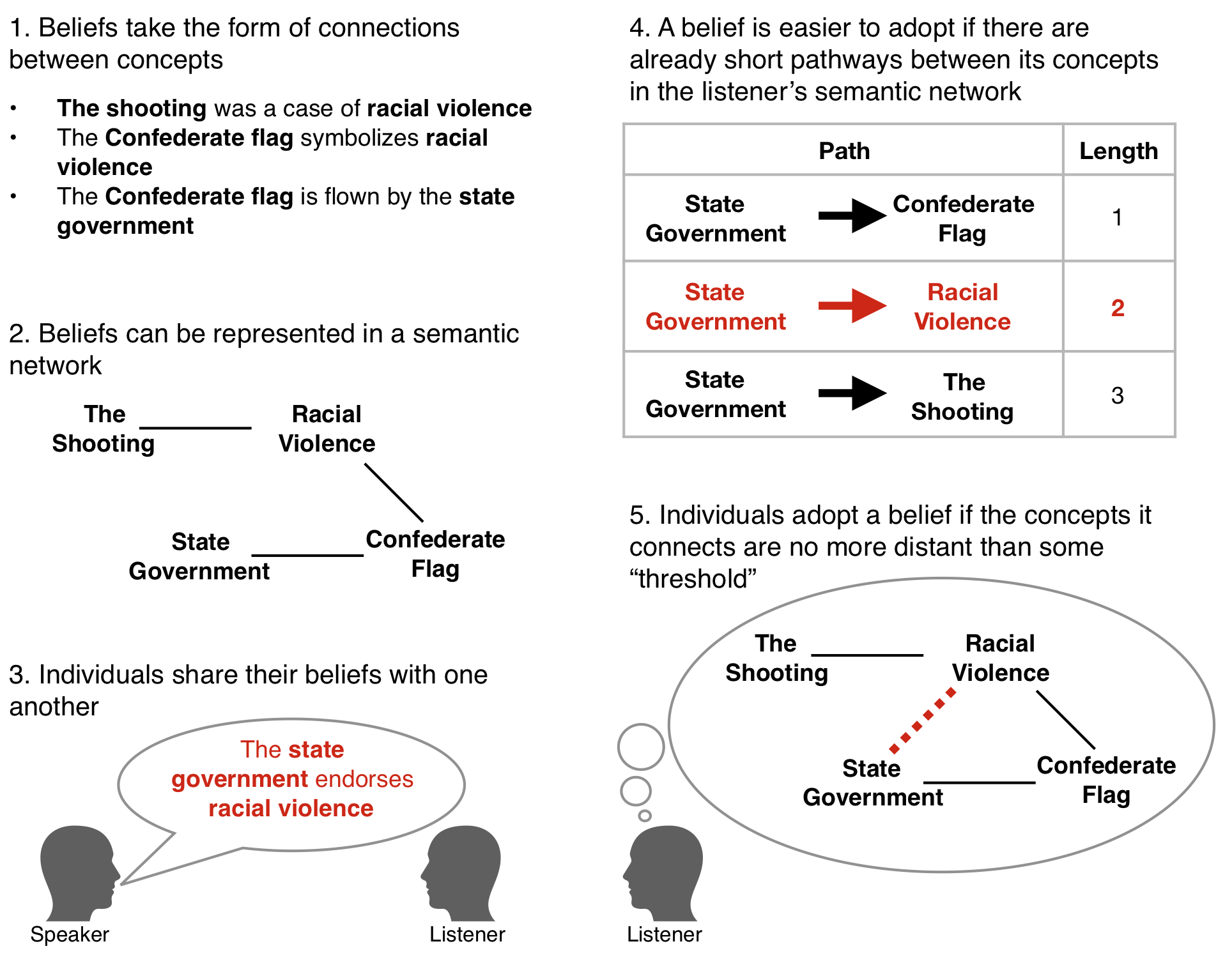
While this paper cannot explore all of the implications of interdependent diffusion, it does highlight a few mechanisms that are of sufficient impact and relevance to make interdependence an unignorable aspect of social contagion. This paper reminds social scientists that the assumption of independence between diffusants is just that – an assumption. Despite its ubiquity, the assumption should be made explicit when employed and challenged when necessary.

**Results:**

*A Minimal Model of Interdependent Diffusion*

A very simple model can illustrate the effect of belief interaction on social contagion. In this model, individuals’ beliefs are represented using the “semantic network” abstraction borrowed from the cognitive science literature *(21-25).* As shown in Fig. 1, nodes in a semantic network represent concepts such as people, places, or activities. A semantic network edge represents the belief that two concepts are connected in some way. This formalization allows beliefs to interact (e.g. beliefs about a place interact via their connection to the node representing that place) without having to pre-specify which beliefs are compatible with one another (e.g. that belief *i* is compatible with *j*, but not *k*)[[3]](#footnote-4).

When an individual is exposed to a new belief by her neighbor in the social network, she decides whether or not to adopt it by seeing how it relates to beliefs in her existing semantic network. She is likely to adopt a belief connecting two concepts that are already close together in her semantic network, as it seems consistent with the beliefs she already holds *(24, 25)*. Conversely, she is unlikely to adopt a belief that two distant concepts are connected, as doing so would dramatically reshape her belief structure. The simplest representation of this tendency is that a simulated individual will adopt any belief that her neighbors possess, as long as the existing distance in her semantic network is below some threshold. To make the simulation easy to follow, I will use a threshold of 2 links distance, such that a belief will be permanently adopted if it closes a triangle in the adopter’s semantic network[[4]](#footnote-5).

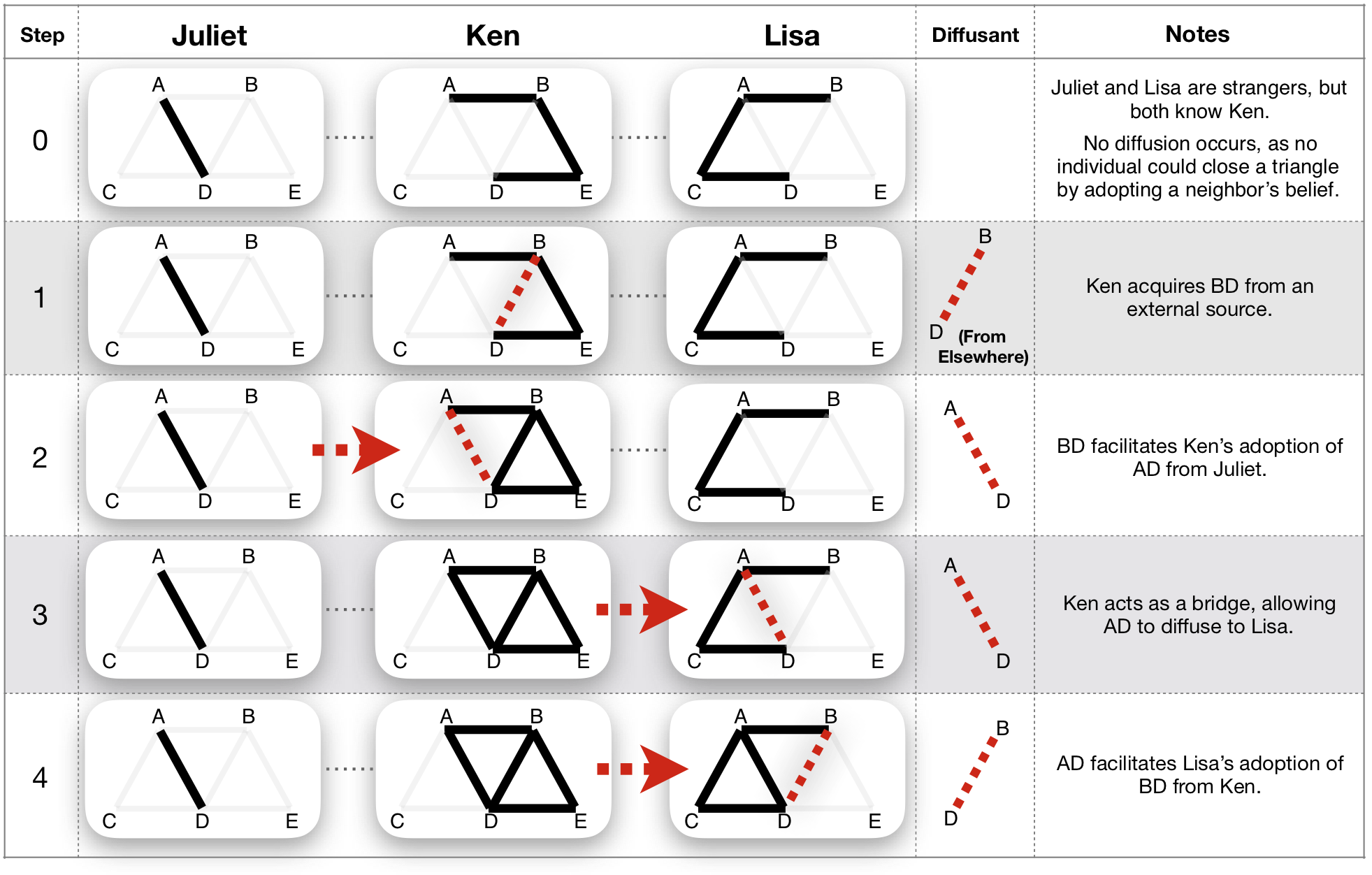


*Fig. 1. A minimal model of interdependent diffusion*

*The Reciprocal Facilitation Mechanism*

To observe the first effect of interdependence on diffusion, we can follow a single focal belief as it spreads through the social network. The focal belief spreads when the prior beliefs of an exposed individual make her susceptible to the focal belief (i.e. in this simplified model the focal belief would “close a triangle” in her semantic network). Continuing to spread to other neighbors in the social network, the focal belief may then reciprocate that facilitation by creating conditions of susceptibility to the beliefs which had previously supported its diffusion, and repeat the cycle.

Fig. 2 illustrates this mechanism of “reciprocal facilitation”, in which simultaneously-diffusing beliefs alternately create susceptibility to one another, and together are adopted by more individuals than any single belief could have reached by diffusing on its own. In this illustration, Juliet, Ken, and Lisa each begin with a different set of beliefs denoted as connections (AB, AC, etc.) between concepts: (A, B, C, etc.) At first, they are not susceptible to adopting beliefs from one another, and so no diffusion can occur. The introduction of a new belief in step 1 kicks off a cycle of reciprocal facilitation in steps 2-4, the introduced belief interacting with the agents’ existing beliefs. In the final state, both the introduced belief and existing beliefs have spread through the network.



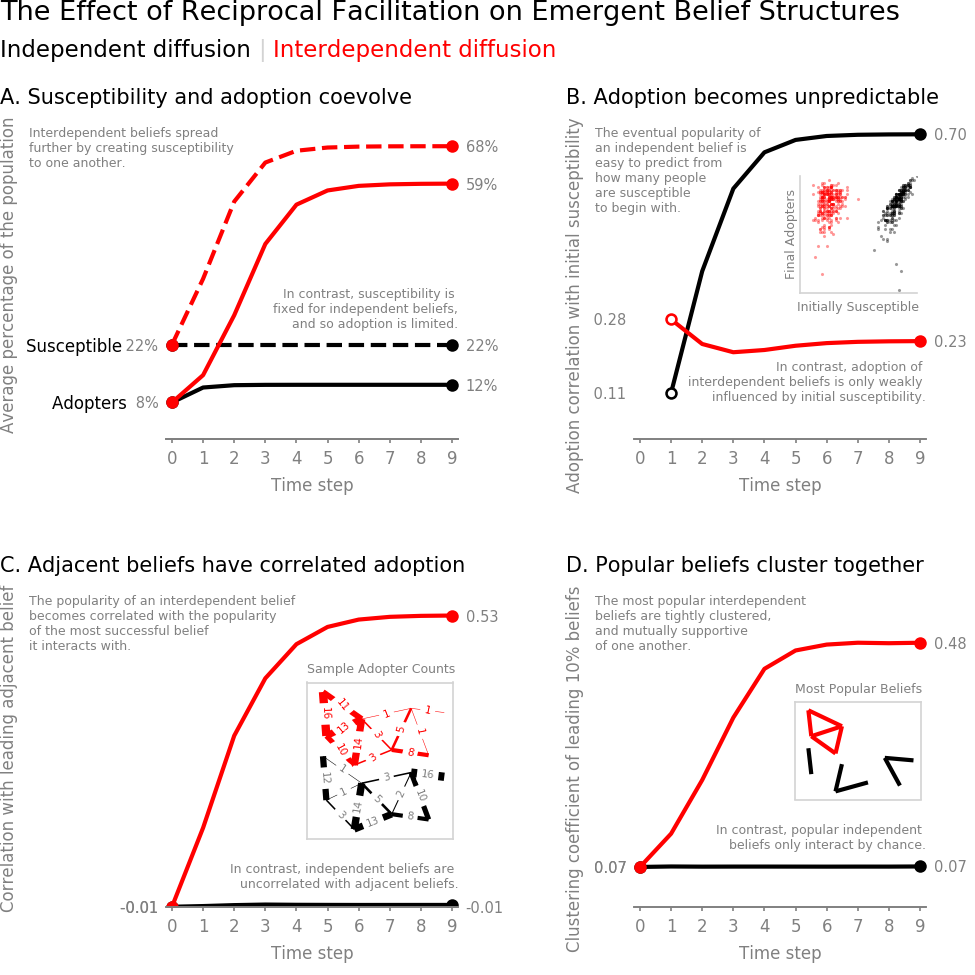
*Fig. 2. The reciprocal facilitation mechanism. Beliefs (AB, AD, etc.) represent connections made between concepts (A, B, C, D, E) and spread when an exposed individual can close a triangle in their existing semantic network by adopting a belief.*

In this example, the diffusion patterns of AB and BD are interdependent. Without the introduction of belief BD, Lisa could never adopt AD, despite her prior susceptibility. Similarly, without AD she could not adopt BD, despite her prior exposure.

What are the effects of reciprocal facilitation at the macro level? Figure 3 shows the result of applying the ‘triangle closing’ decision rule to a population of 60 agents in a random social network. For details of the simulation, see the “Methods” section. The first effect of reciprocal facilitation is to allow the average number of susceptible and adopting individuals to grow simultaneously, as shown by the red logistic growth curves in Fig. 3A. As a comparison, in black, I show how the same beliefs would diffuse if they did not interact. In the independent case, beliefs may only be adopted by individuals who are susceptible at the start, and so beliefs spread less widely in the population. Put another way: to explain the same level of final adoption, models of independent diffusion must assume significantly more initial susceptibility to each belief.

It would be reasonable for us to ignore interdependence and instead assume more widespread initial susceptibility, were it not for the second effect of reciprocal facilitation. When diffusants do not interact, the population initially susceptible to a belief is an excellent predictor of the number who will eventually adopt it. However, Fig 3B shows that as susceptibility evolves alongside adoption in interdependent diffusion, the population that is initially susceptible to a belief becomes a poor predictor of who will eventually adopt it.

To use a concrete example, imagine that a “focus group” is selected from our artificial population before the simulation starts, and imagine that amongst this group AB is adopted by 25% more people than CD. Unsurprisingly, when we simulate independent diffusion in the full population, AB turns out to be more popular than CD in over 99% of cases. On the other hand, when we simulate interdependent diffusion, AB is adopted by more people than CD only 57% of the time – just slightly better than chance.



*Fig. 3: The effect of reciprocal facilitation on emergent belief structures.*

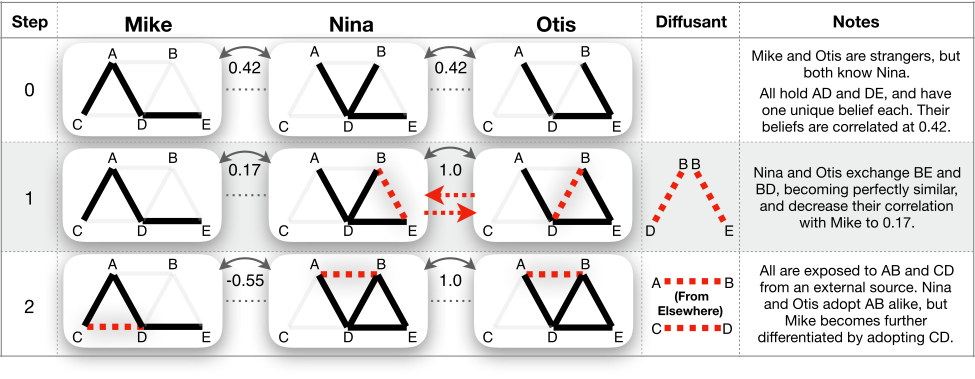
The popularity of interdependent beliefs is hard to predict because reciprocal facilitation does not support the diffusion of all beliefs to the same degree. Instead, it preferentially amplifies those that are connected to other widely adopted beliefs. Figure 3C shows that a belief’s popularity becomes correlated with the popularity of the most popular belief adjacent to it. This is because a belief that is supported by many popular beliefs will have many opportunities to diffuse, and then to facilitate the adoption of other closely related beliefs. Conversely, a belief that is connected only to unpopular supporting beliefs will have trouble reaching even the few individuals who are susceptible to adopting it.

As a result, patterns emerge when individual semantic networks are aggregated to the level of the population. After interdependent diffusion, the most popular beliefs are closely related to one another and form densely connected islands in the semantic space. Fig. 3D shows one of many possible measures of this structure, the clustering coefficient (see the methods section for others). While clustering is a consistent outcome of interdependent diffusion, the process itself is strongly path-dependent, and so the precise locations of the clusters are unpredictable. This simulation suggests that clustering may occur not because some beliefs are inherently more compatible with one another, but because the most popular beliefs bring popularity to their neighbors. In contrast, independently diffusing beliefs can only exhibit this kind of macro-level structure if it is assumed to be present in individuals’ initial susceptibility to beliefs. However, assuming so merely defers explanation to another exogenous phenomenon.

*The Agreement Cascade Mechanism*

To explain a second mechanism of interdependent diffusion, we can observe a pair of neighbors in the social network. When these two individuals exchange beliefs, they become more similar to one another. Because their existing belief sets influence the way they respond to new beliefs, shared beliefs make the two individuals more likely to adopt (or reject) the same new beliefs in the future, regardless of any preference to align or distinguish themselves from one another. As a result, they become more similar still. Additionally, similar neighbors expand one another’s access to beliefs that the two may adopt in common, and filter each other’s exposure to beliefs that would set them apart from one another. In contrast, individuals who adopt differing beliefs are likely to diverge further as their dissimilar semantic networks make them susceptible to different beliefs.

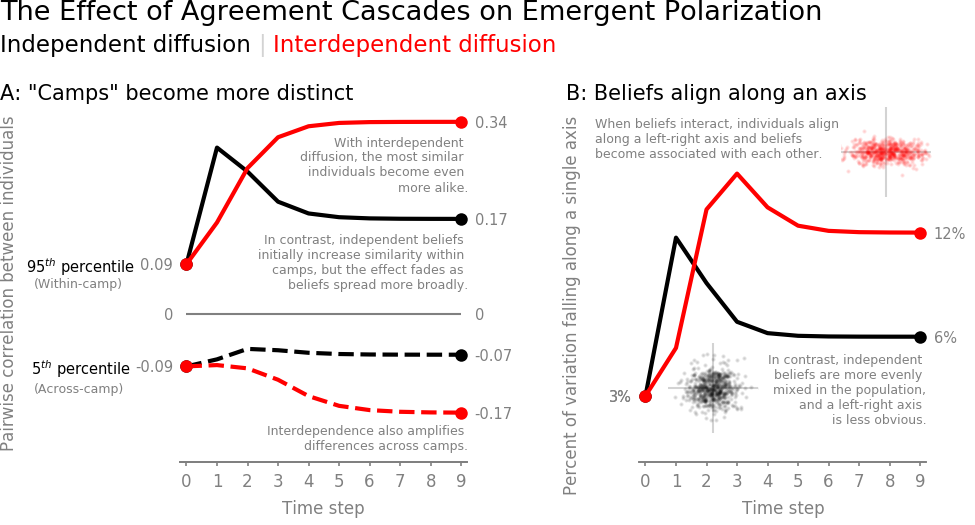
Fig. 4 illustrates this “agreement cascade” mechanism, in which exchanging beliefs creates further opportunities for similarity to develop between individuals. In this example, Mike, Nina, and Otis share two out of their three beliefs with their neighbors (AD and DE), and so their belief sets are positively correlated at 0.42. When Nina and Otis exchange beliefs in step 1, they not only make themselves more similar in the act, but they also set themselves up for congruent behavior in the future. Their response to externally supplied beliefs in step 2 further reinforces their similarity with one another and amplifies their emerging dissimilarity with Mike.



*Fig. 4. The agreement cascade mechanism. Similarity between neighbors is measured as the correlation (phi coefficient) between the set of beliefs they hold.*

Returning to our 60-agent simulation, we can ask how patterns of similarity evolve at the macro level. In this simulation, we assume that adoption is permanent, and so diffusion raises the average similarity of the population purely by chance. This is especially true in the first round of simulation when neighbors have had a chance to exchange beliefs, but those beliefs have not had an opportunity to spread through the whole population. Fig. 5A shows that agreement cascades drive individuals who already share some beliefs to form increasingly self-similar “camps”, and amplify the differences between camps.

A simple and reproducible way to assess the similarity of individuals within an ideological camp – absent exogenous labels such as demographic or party – is to measure the similarity between all pairs of individuals and define a certain percentile as belonging to the same ideological camp. The more exclusive we are (i.e. the higher the percentile), the more conservative the claim that these represent “within-camp” relationships. Similarly, to define across-camp similarity, we can define a percentile that (conservatively) represents relationships between individuals in different ideological camps. In this simulation, I use the 95th and 5th percentiles respectively. See the methods section for more details[[5]](#footnote-6).

**

*Figure 5: The effect of agreement cascades on emergent polarization*

As individuals become more similar to people within their own camp and more differentiated from people in other camps, correlation begins to emerge between belief sets. For example, AB may co-occur with CD 70% of the time, but co-occur with DE only 10% of the time. This sorting is a form of dimensionality reduction, in which the complex landscape of beliefs becomes condensed into a few axes (e.g. liberal-conservative, or libertarian-populist) *(26)*. The upper inset diagram in Fig. 5B illustrates individuals who exchange interdependent beliefs and begin to take positions along a left-right axis. Differences between these individuals are mostly due to their relative positions along the left-right axis, and individuals at any point on the axis are relatively similar to one another. The lower inset diagram illustrates individuals who have exchanged independent beliefs with one another. In this diagram, there is less alignment along the left-right axis and more variation between individuals at any given point.

Fig. 5B measures belief alignment as the amount of variation between individuals that can be explained by the best fitting axis in the space of possible beliefs (see the methods section for details). While some difference from the homogeneous initial state is an inevitable result of diffusion, the effect of agreement cascades is to increase the ease with which individuals’ beliefs can be summarized by their position on a “political spectrum”. As a result, it becomes easier to predict an individual’s position on one belief from their position on other beliefs.

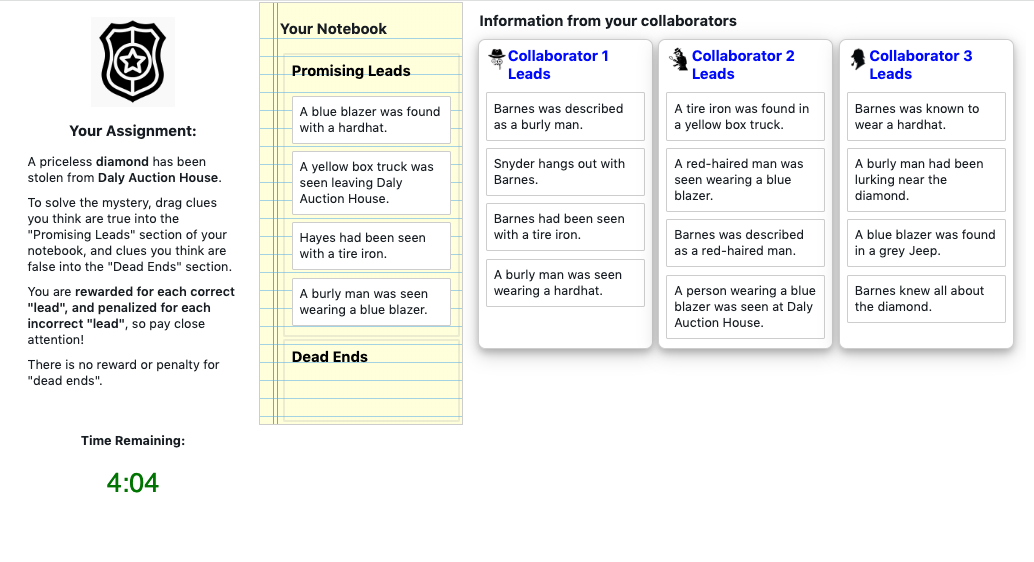
*A Laboratory Experiment*

The above simulation creates a strong theoretical prior with specific, testable predictions about the effect of interdependence on diffusion. However, all such simulations deliberately simplify human behavior to explain the sociological mechanisms in play. To have confidence in the proposed theory, we must empirically test its predictions with human actors and realistic beliefs. Unfortunately, this is difficult to do in the wild, where demographics, social network dynamics, and existing belief systems also influence the outcome, and where independence between treatment groups is difficult to establish.

To overcome these challenges, I designed an online laboratory experiment in which individuals were randomized to positions in an artificial social network, blinded as to one another’s identities, and given a constructed set of beliefs to exchange. A total of 2,768 U.S. and Canadian workers were recruited from Amazon Mechanical Turk, of which 2,400 participants completed training and were randomized into 120 separate 20-person social networks[[6]](#footnote-7). Each network was assigned one of four experimental conditions, and matched comparisons were made with n=30 samples in each condition. The participant population was 45% female; mean 37.1 years old; 27% high-school, 49% bachelors, 16% masters+ graduates. 96.8% of players who completed training went on to complete all steps of the experiment, with less than 0.4% difference in dropout between conditions. The preregistration for this experiment[[7]](#footnote-8) included all code necessary to generate test conditions, implement the game using Empirica *(33)*, conduct the experiment, and process the resulting data.

In this experiment, participants were asked to find a solution to a mystery by identifying a burglar's name, description, clothing, burglary tool, and getaway vehicle. Participants were seeded with four clues to the mystery, and incentivized to sort those clues into "Promising Leads" and "Dead Ends", as shown in Fig. 6. When a participant categorized a clue as a Promising Lead, it was immediately shared with their three neighbors in the 20-person social network.

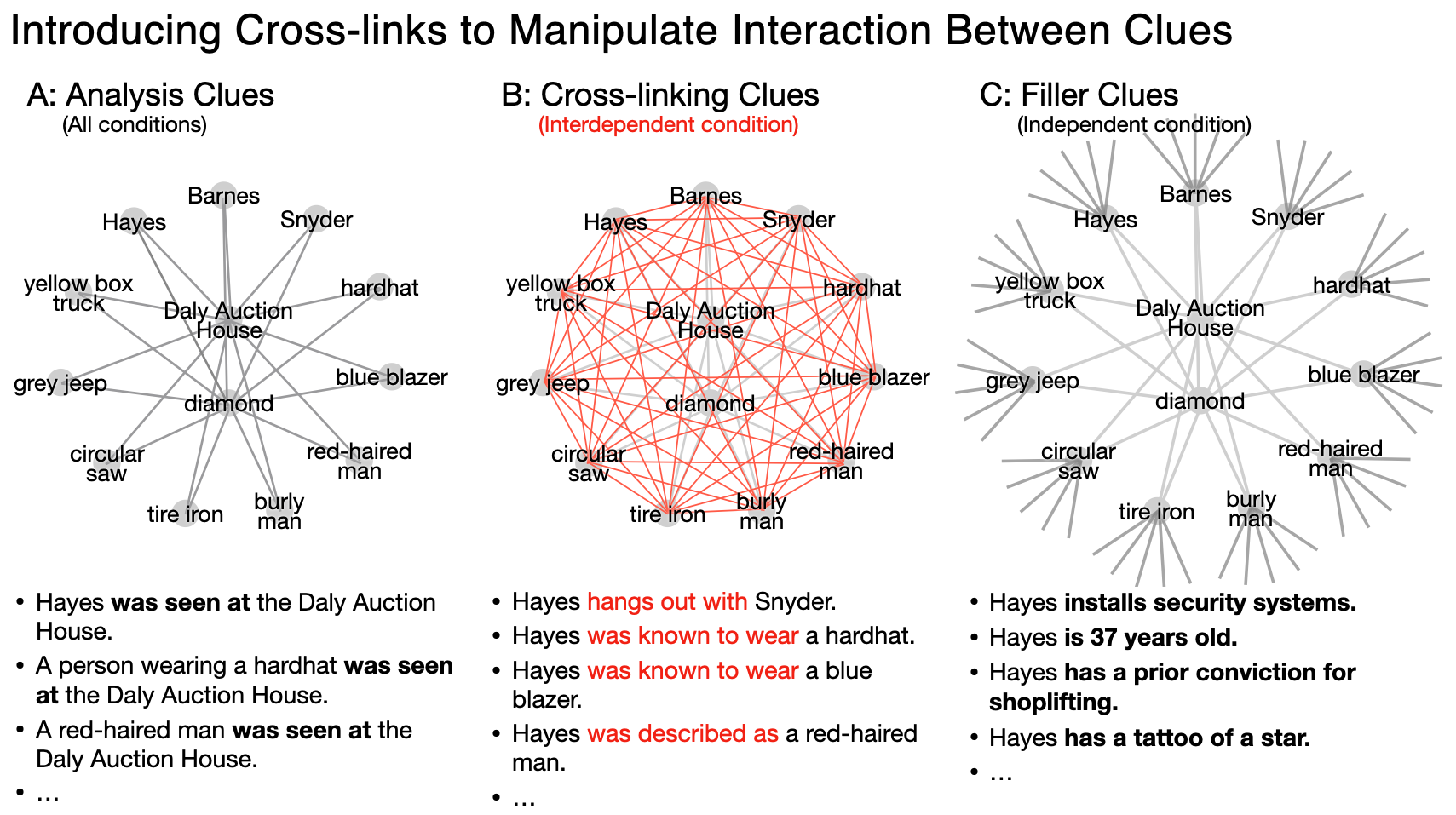
The measures of polarization described in the above simulation were operationalized to reflect participants’ behavior, and also their self-reported opinions. “Behavioral” measures were constructed from each player’s final categorization of 22 clues, reflecting the cumulative choices that individuals had made throughout the game. Following the game, participants estimated the likelihood that each suspect, vehicle, etc. was involved in the crime. “Self-report” measures were constructed from these 11 post-game assessments, reflecting how individuals internalized the information they encountered to create opinions. Finally, participants were asked to rate their confidence in their overall solution, and the level of consensus they perceived in their team. The “Gameplay” section of the supplement recounts the participants’ full experience.



*Fig. 6. The primary user interface of the “Detective Game”. Clue cards can be dragged into categories in the player’s “Notebook”. Promising Leads are immediately shared with 3 neighbors, who can drag them into their own notebooks.*

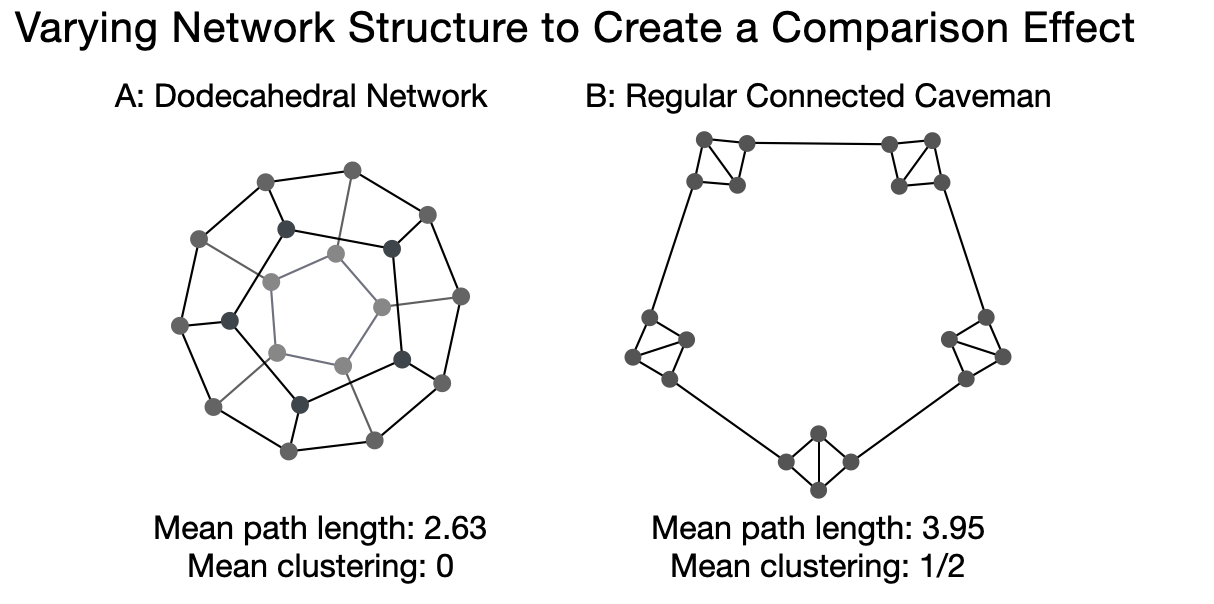
In the theory-building simulations above, agents can be programmed to pay attention to interactions between beliefs in an “interdependent” world, and ignore those interactions in an identical “independent” world. Human beings, on the other hand, are wired to see connections and relationships between ideas, and so it is impossible to conduct a perfect test in which *all* clues interact in one experimental condition, while *none* of those clues interact in another. Instead, I partitioned the clues such that some were common to both independent and interdependent conditions and used for analysis (i.e. “analysis clues”), while others varied across conditions to manipulate the level of interaction between analysis clues. In each game, 22 analysis clues linked the crime scene and stolen object to each of the suspects, descriptions, clothes, tools, and vehicles, as illustrated in Fig. 7a. In the interdependent condition, 55 additional “cross-linking clues” connected all of these elements of the mystery (i.e. suspects, vehicles, etc.) to one another (Fig. 7b). In the independent condition, 55 “filler clues” took the place of cross-linking clues to break the relationships between elements (Fig. 7c). These filler clues ensured that the two conditions had the same total number of clues, each as similar to the other as possible.

Clues were extensively pre-tested to minimize bias from the outside world, such that each element was perceived to be equally likely to be involved in a burglary absent other information. Sets of clues were randomly generated from the pool of pretested concepts, representing over 18 million possible mysteries. At the start of the game, each clue was present in exactly one player’s notebook. As a result, beliefs spread on a level playing field, such that *a priori* none should be expected to diffuse more than any other. Further information can be found in the “Clue Generation” section of the supplement.



*Fig. 7: Clues are designed such that in the interdependent condition, “analysis” clues are connected to one another by cross-linking clues; while in the independent condition, analysis clues remain disconnected from one another.*

In addition to the two types of clue structures, the experiment also included two types of social networks, in both of which 20 players were connected to exactly three neighbors each. The first network was a “dodecahedral” network (Fig. 8a), in which none of a player’s neighbors were directly connected to any other, and the average network distance between individuals was short. We should expect to find very little polarization in this network, as information can diffuse across the network readily, and coordination among subgroups is impeded by the lack of mutual connections. The second social network was a “regular connected caveman” structure (Fig. 8b), which exhibits high levels of clustering and large average distances between individuals. We should expect to find high levels of polarization in this network regardless of the level of interaction between clues, as strong clustering makes it easy for subgroups to converge on a shared set of clues, and long average path lengths make it harder for information to spread between camps. Further information can be found in the “Social Network Structure” section of the supplement.



*Fig 8. Participants inhabit 20-node regular degree-3 networks chosen to (A) minimize characteristic path length and clustering, and (B) maximize these attributes.*

Together these manipulations create four separate conditions. The dodecahedral network and independent clue set formed a baseline condition. Contrasted with the baseline, experimental conditions measured the effect of interdependence, the effect of social network structure, and the effect of both manipulations combined. Blocks of four simultaneous games were constructed with one game in each condition. To guard against latent external biases, each block used a different randomly-generated set of clues. Clue assignments varied as little as possible within each block, such that the clues assigned to a particular network position in one game corresponded to those assigned to the same network position in the other three games. Upon completing training, participants were randomly assigned to positions within a block, blind to their treatment condition. I treated games within each block as matched samples, using one-sided pairwise t-tests to assess the differences between each experimental condition and the baseline.

Experimental Results

Over eight minutes of gameplay, participants on average made 28.4 classifications and adopted 16.2 clues as promising leads. Even though there were no solutions to the given mysteries, and clues were perfectly symmetric and seeded equally in the population, participants came to strongly-held beliefs about which suspect was guilty and how they performed the crime. For example, over half of participants reported confidence in at least one aspect of the mystery of 95% or greater, and the average difference in confidence between the most and least likely suspect, vehicle, tool, etc. was over 30% in each category.

The baseline condition consisted of independent clues diffusing through the non-polarizing dodecahedral social network. This condition approximates the conditions of independent diffusion assumed broadly in the social contagion literature. However, despite the level of control afforded by a laboratory experiment, this condition does not perfectly eliminate all interactions between clues. For example, even though the independent clue set contains no explicit links between suspects, participants may draw an implicit connection between the guilt of one suspect and the presumed innocence of another. This imperfect control means that measured differences between independent and interdependent conditions will underestimate the true effect of belief interaction.

*Finding 1: Interdependent diffusion has a measurable polarizing influence*

The first finding of this experiment confirms the theoretical prediction that interdependence between diffusants has a previously unobserved polarizing influence on the population’s adoption behavior. Increasing the level of interdependence between clues above the baseline condition (all else held constant) measurably increased the population’s alignment along a left-right axis among both behavioral and self-reported measures of belief (Table 1 and Fig. 9). Although not all similarity measures were significant, experimental results support the prediction that camps became more self-similar and more distinct from one another as a result of interdependence.

The above theory does not predict how individuals will feel about their team’s performance. Interdependence did not change the perceived consensus among the team by any measurable amount, despite the increase in polarization. However, participants in the interdependent condition did report feeling more confident in their solution.

*Finding 2: The effect of interdependence is of meaningful size*

The second finding is that the measured effect of interdependence is large enough to be worth attention when compared to other drivers of polarization. While we should be wary of generalizing macro-scale effect sizes from lab experiments to large-scale social networks, we can meaningfully compare the effect of interdependent diffusion with the more familiar effect of social network structure. An *a fortiori* test is to compare the measured effect of interdependence with the maximal effect that can be induced by changing the social network structure alone. If the effect of interdependence is relevant compared to that of drastic changes to the network structure, then it should also be relevant when compared with more moderate effects of network structure in the real world.

The two networks included in this experiment exhibit dramatically different levels of network clustering and average path length. However, regardless of the strength of the social network manipulation, the effect of interdependent diffusion was only significantly different from the effect of social network structure in two of the six outcome measures: the behavioral measures of alignment along a left-right axis and within-camp similarity. At the same time, the estimated effect of interdependence reached approximately 40% of the *a fortiori* social network effect for the behavioral measures of across-camp similarity and alignment along a left-right axis. Other measures reported smaller but still meaningful fractions (Table 1 and Fig. 9).

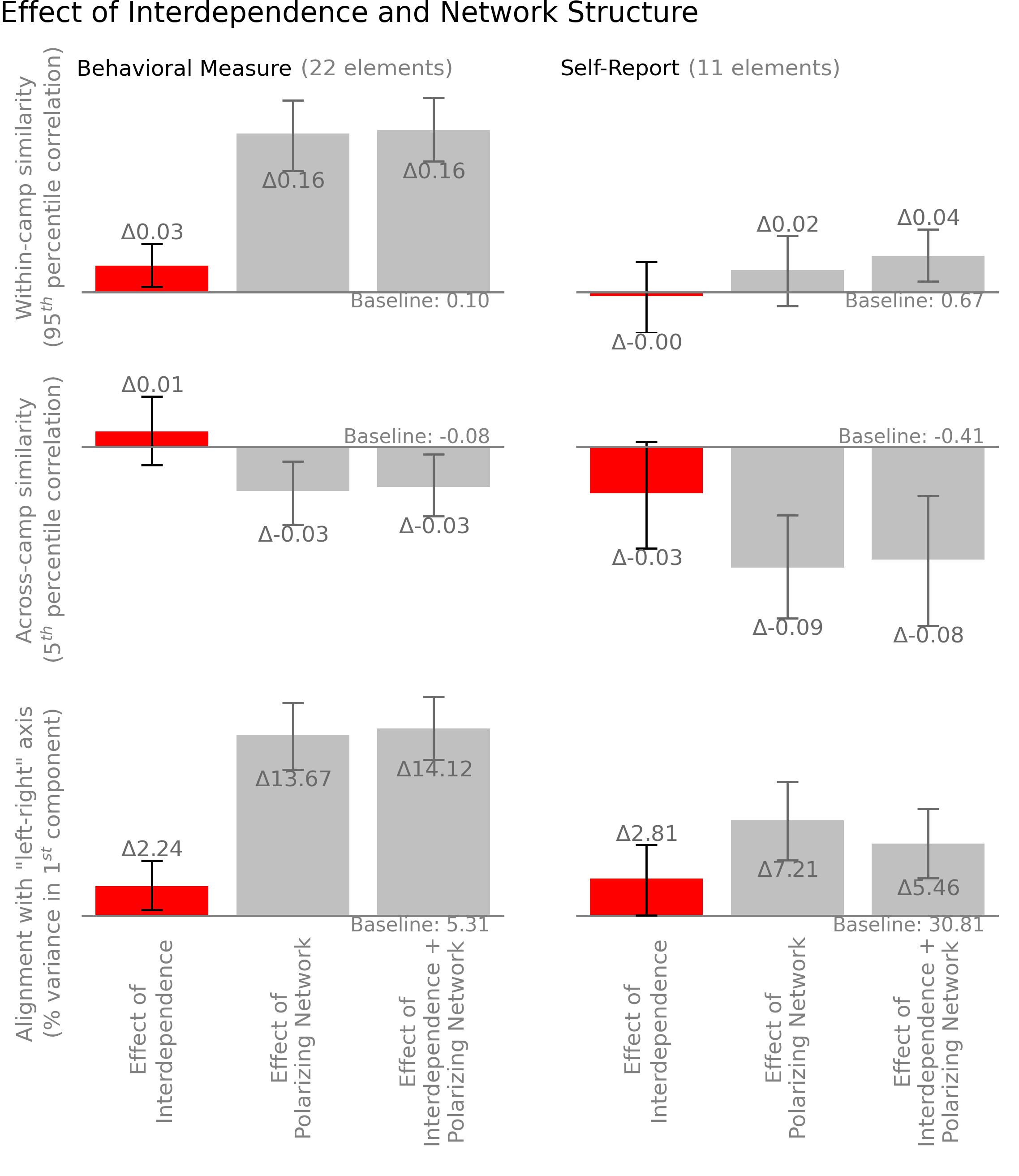
Again, these should not be taken as estimates of the effect size we would expect to see outside of the lab. However, the fact that the effects of interdependence and social network structure were even comparable in this experiment suggests that scholars’ almost exclusive focus on network structure over interdependence is out of proportion to the relative importance of the two effects.

*Finding 3: A negative interaction effect likely exists between interdependent diffusion and social network structure*

A practical question is whether and how polarization can be reduced. Shortening the average distance between individuals[[8]](#footnote-9) and breaking up clusters seems a plausible strategy for helping information travel across the network before camps consolidate. If there is no interaction (or a positive interaction) between the effects of interdependence and network structure, changes to social network structure would be at least as powerful when beliefs interdepend as they are when beliefs spread on their own. However, if there is a negative interaction between interdependence and network structure, then models that assume interdependence between diffusants will overestimate the gains that can be made by changing the network structure alone.

The final condition of this experiment measured the effect of interdependent diffusion in the polarizing social network. While statistical power was poor, this condition revealed a significant negative interaction effect for alignment with a political axis based on self-reported beliefs. It is unclear whether the interaction arises because the two factors operate through similar pathways, or because the outcome variable begins to saturate. However, it is clear that even if we had perfect control of social network structure, we could not fully address polarization without attending to belief interaction.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Table 1. Effect and interaction summary** | | | | | | |
|  |  | **Effect over baseline condition** | | | | Interaction of  Interdependence and Network Effect |
|  |  | Interdependent Clues Alone | Polarizing Network Alone | Interdependent Clues and Polarizing Network | |
| Within-Camp Similarity *(95th percentile correlation)* | Behavioral1 | +0.0249\*\* | +0.155\*\*\* | | +0.158\*\*\* | -.0227 |
| Self-Report1 | -0.00389 | +.0231 | | +.0357\*\*\* | +.0183 |
| Across-Camp Similarity *(5th percentile correlation)* | Behavioral1 | +0.0112 | -.0307\*\*\* | | -.0310\*\*\* | -.00827 |
| Self-Report1 | -0.0345\* | -.0904\*\*\* | | -.0844\*\*\* | +.0402 |
| Alignment with Political Axis *(% Variance in 1st Component)* | Behavioral1 | +2.16%\*\* | +13.8%\*\*\* | | +14.2%\*\*\* | -1.79% |
| Self-Report1 | +2.77%\*\* | +7.16%\*\*\* | | +5.34%\*\*\* | -4.57%\*\* |
| Confidence | Self-Report | +2.17%\*\* | +0.682% | | +3.74%\*\* | +.887% |
| Consensus | Self-Report | -0.624% | +1.75%\* | | +1.03% | -.0925% |
| 1 Preregistered; \*P value <.1, \*\*P value <.05, \*\*\*P value <.01; one-sided pairwise T-tests; n=30 pairs; Baseline: Independent clues and non-polarizing social network. | | | | | | |



*Fig. 9. Interdependent diffusion has a measurable polarizing effect over a baseline condition of independent clues in a non-polarizing social network. The effect size is an unignorable fraction of that of the* a fortiori *social network manipulation. A negative interaction effect is likely. (Error bars show 95% CI).*

**Discussion:**

When multiple interrelated beliefs are regularly found together *(1, 15)*, it is easy to suppose that they constitute internally-consistent worldviews that represent some underlying truth about the world. Through both simulation and experiment, this paper demonstrates that external “truth” need not be present for these worldviews to emerge. Individuals may observe internal consistency between their beliefs and social confirmation of their choices entirely as a product of interdependent diffusion. It is perhaps alarming that this can be observed in an experiment premised on a *designed* lack of evidence for any conclusion, in which participants interacted with the presented information for only *eight minutes*.

Implicit in the above supposition is the idea that observed worldviews represent natural groupings of beliefs (right or wrong), and also that other groupings could not have become popular. Instead, this paper shows that worldviews can emerge in a path-dependent way that varies due to chance encounters and adoptions.

It is not surprising to see polarization develop when a social network is formed from relatively isolated communities *(8, 15)*, or to see polarization grow when individuals choose to associate with people like themselves *(27)*. What is surprising is that when beliefs interact, polarization emerges even in a static, symmetric, and well-connected social network whose members are anonymous, have no desire to form groups, have no predefined issues to disagree over, have no incentive to ‘perform’ for their neighbors, and indeed have very little emotional connection with the beliefs they are sharing. It turns out that none of our habitual explanations for polarization are strictly necessary; interdependent diffusion is sufficient.

Of the many difficulties in designing a naturalistic experiment to manipulate interdependence between otherwise balanced diffusants, the most complex was to create and test a nontrivial set of “independent” beliefs for participants to exchange. Despite many rounds of structured pretesting with hundreds of participants, in the end, the “control” condition of this experiment only approximates conditions of independence for a subset of the diffusants. The real-world relevance of interdependent diffusion could not be better emphasized than by the difficulty of recreating the almost-ubiquitous assumption of independence between diffusants *even in ideal laboratory conditions*. It may well be that rather than assuming independence between diffusants, we ought to consider simultaneously spreading beliefs to be interdependent until proven otherwise. It is certainly the case that when scholars assume independence between diffusants, that assumption should be noted and challenged.

While this paper demonstrates the relevance of belief interaction for polarization and worldview emergence, it is by no means an exhaustive study of the consequences of interdependent diffusion. When the assumption of independence is relaxed in the broader literature, further effects will certainly be discovered and their practical implications realized. Rather than closing the book, I hope that this paper opens a new chapter in the study of interdependent diffusion.

Several weeks after the Charleston Shooting, the South Carolina state legislature finally approved a bill to remove the Confederate flag from the State House grounds. Why was the Charleston shooting the incident that spurred this change, while other instances of racial violence throughout history had not had the same effect? The removal of the flag was not a new issue, nor an inevitable outcome *(1, 28)*. The legislators did not adopt the belief that the flag should be removed because they were newly exposed to it, but because of a series of facilitating beliefs regarding the flag, racial violence, symbols of oppression, and the beliefs of their constituency. These influences cannot be reduced to simple contagion of independent beliefs, but depend entirely upon the interaction of interdependent, simultaneously diffusing beliefs.

**Methods:**

In the simulation presented in Figs. 3 and 5, the social network is a connected Erdős–Rényi (Gnm)random graph with 60 agents total, each with an average of 3 neighbors. Each agent is initialized with 25 beliefs (edges) selected randomly from the 300 edges available in a complete semantic network with 25 concepts (nodes). These values ensure good coverage of beliefs in the population, while individual semantic networks are initially sparse. Random seeding ensures that the simulation starts without polarization or systematic variation in belief popularity, and also that the social network structure itself does not contribute to polarization. Because beliefs are drawn from a complete semantic network, there is no natural belief structure around which polarization may nucleate.

In each step, individuals are selected in random order, and update their beliefs by incorporating into their semantic networks all beliefs (edges) that their neighbors possess and they are susceptible to adopting. In the interdependent case, individuals are susceptible to any belief that closes a triangle in their semantic networks at the current time. In the comparison (independent) case for Fig 3A, a random selection of the population is defined to be susceptible to each belief in the same proportion as are initially susceptible to the belief in the interdependent diffusion case. In Figs 3B-D and Fig 5, a random selection of susceptible individuals is made in proportion to the *final* number of susceptible individuals in the interdependent diffusion case. As a result, a histogram of the extent of diffusion of each belief is approximately the same under both independent and interdependent treatments. This ensures that the subsequent presentation of results reflects purely the effect of interdependence between diffusants, and not the effect of different levels of adoption in the compared populations. Results are similar when calculated based upon the initial susceptibility.

The measures presented Figs. 3A-C are averaged over all beliefs in the simulation, and over 20,000 simulations. The measures in Figs. 3D and 5 are population-level measures averaged over 20,000 simulations. This volume of simulations is an order of magnitude beyond the point at which noise affects the result. The measure of the susceptible population in Fig. 3A (and all discussion of the susceptible population) represents all individuals who would adopt the belief if exposed to it, along with all of the individuals who have already adopted the belief. I break from the traditional compartmental model description of susceptibility and adoption/infection as mutually exclusive categories to more effectively demonstrate the coevolution of susceptibility and adoption. Using traditional definitions does not change the outcome of the simulation. Fig. 3B measures the correlation between new adoption and those who are initially susceptible to the belief but do not start with it. As this has no meaningful value at t0, the curve is drawn from t1-t9. Figure 3D uses the clustering coefficient of a semantic network constructed from the most popular 10% of beliefs as a demonstration that the most popular beliefs are mutually interrelated, and not merely all related to a single leading belief (e.g. a star or barbell pattern). Clustering only makes sense when beliefs are conceptualized as a semantic network. Other conceptualizations of belief interaction might prefer to plot the number of top decile beliefs that each top decile belief interacts with. This measure gives essentially the same result (i.e. large fractional growth over time in the interdependent case, with no change from randomness in the independent case) but lacks the depth of meaning captured by the clustering metric. The measure is generally insensitive to the specific threshold used to define a ‘popular’ belief for any thresholds between about 5% and 40%. Fig. 5B shows the percentage of variation explained by the first principal component of the data. The original feature space has one dimension for each belief in the simulation (300), and points representing each individual’s position in that feature space (60) according to the beliefs they have adopted. The inset graphs are exaggerated and show a larger population to illustrate how a component can explain more or less variation. For all remaining parameters in the simulation, other (meaningful) values and measures yield qualitatively similar results. For a full sensitivity analysis and code, see <https://github.com/JamesPHoughton/interdependent-diffusion>.

There are many ways to measure the similarity between individuals. The “self-reported” beliefs of experiment participants fall on a continuous scale from 0 to 100, and so it is natural to use Pearson’s correlation on the vectors of individuals’ beliefs. This measure has the advantage of being easily interpretable and having a well-defined range that is independent of the number of features in the vector of attributes being compared, and the negative region of which can be interpreted as expressing dissimilarity. To assess the similarity of the binary “behavioral” data I use the Phi coefficient, an analogous measure to Pearson’s correlation with the same interpretable range. The behavioral measures in the experiment are sensitive not only to interdependence and network structure but also to the average level of diffusion of beliefs. To minimize noise due to differences in the level of activity between games, each of the behavioral measures is assessed compared to what would be expected due to chance, keeping the number of adopters of each clue and the number of clues adopted by each participant fixed. All code for conducting the experiment and analyzing results, along with the results of the experiment, are available at https://github.com/JamesPHoughton/detective-game-interdependent-diffusion.

References:

1. J. P. Houghton *et al.*, Beyond Keywords: Tracking the Evolution of Conversational Clusters in Social Media. *Sociol. Methods Res.* **48**, 588-607 (2019).
2. M. Granovetter, The Strength of Weak Ties. *Am. J Sociol.* **78**, 1360-1380 (1973).
3. M. Granovetter, Threshold Models of Collective Behavior. *Am. J Sociol*. **83**, 1420-1443 (1978).
4. T. Schelling, “Sorting and Mixing” in *Micromotives and Macrobehavior*. (Norton, Toronto, 1978) chap. 4.
5. D. J. Watts, A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci. U.S.A.* **99** 5766-5771 (2002).
6. R. Burt. Structural Holes: the social structure of competition. *Harvard University Press* (1992).
7. D. Kempe, J. Kleinberg, and E. Tardos. Maximizing the Spread of Influence through a Social Network. *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2003.
8. D. Centola, M. Macy, Complex Contagions and the Weakness of Long Ties. *Am. J Sociol.* **113** 702-734 (2007).
9. J. Travers and S. Milgram. An Experimental Study of the Small World Problem. *Sociometry.* **32** 425-443 (1969).
10. J. Lorenz et al. How social influence can undermine the wisdom of crowd effect. *PNAS* **108** ﻿9020-9025 (2011).
11. L. Muchnik, S. Aral, and S. Taylor. Social Influence Bias: A Randomized Experiment. *Science* **341** 647-651 (2013).
12. D. Rand, S. Arbesman, and N. Christakis. Dynamic Social Networks Promote Cooperation in Experiments with Humans. *PNAS* **108 ﻿**19193-19198(2011)
13. S. Suri and D. Watts. Cooperation and Contagion in Web-Based, Networked Public Goods Experiments. *PLoS ONE* **6** ﻿e16836 (2011)
14. D. Baldassarri, P. Bearman, Dynamics of political polarization. *Am. Sociol. Rev.* **72** 784-811 (2007).
15. D. DellaPosta *et al.*, Why do liberals drink lattes? *AM. J. Sociol.* **120** 1473-1511 (2015).
16. N. E. Friedkin *et al.*, Network science on belief system dynamics under logic constraints. *Science.* **345**, 321-326 (2016).
17. A. Goldberg, S. Stein, Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation. *Am. Sociol. Rev.* **83** 897-932 (2018).
18. C. T. Butts, Why I know but don’t believe. *Science.* **354** 286-287 (2016).
19. S. E. Parsegov *et al.*, Novel Multidimensional Models of Opinion Dynamics in Social Networks. IEEE Trans. Automat. Contr. **62** 2270-2285 (2017).
20. F. Xiong *et al.* Analysis and application of opinion model with multiple topic interactions. *Chaos.* **27** 083113 (2017).
21. A. M. Collins, E. F. Loftus, A Spreading-Activation Theory of Semantic Processing. *Psychol. Rev.* **82** 407-438 (1975).
22. M. Steyvers, J. B. Tenenbaum, The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cogn. Sci.* **29** 41-78 (2005).
23. R. J. Brachman, “On the epistemological status of semantic networks” in *Associative Networks.* 3-50 (Acad. Press, 1979).
24. M. Schilling, A ‘small-world’ network model of cognitive insight. *Creat. Res. J.* **17** 131-154 (2005).
25. Nickerson, Raymond S. Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. *Review of General Psychology* **2** 175–220 (1998).
26. P. DiMaggio, J. Evans, B. Bryson. Have Americans Social Attitudes Become More Polarized? *Am. J Sociol.* **102** 690-755 (1996).
27. P. Dandekar, A. Goel, and D. Lee. Biased assimilation, homophily, and the dynamics of polarization. *PNAS* **110** 5791-5796.
28. J. Becker, E. Porter, and D. Centola. The wisdom of partisan crowds. *Proc National Acad Sc*i. **116**, 10717–10722 (2019).
29. Permanyer, I. The conceptualization and measurement of social polarization. J Econ Inequal 10, 45–74 (2012).
30. Baldassarri, D. & Gelman, A. Partisans Without Constraint: Political Polarization and Trends in American Public Opinion. *Ssrn Electron J* **114**, 408–446 (2008).
31. Poole, K. T. & Rosenthal, H. The Polarization of American Politics. *J Politics***46,** 1061–1079 (1984).
32. Esteban, J.-M. & Ray, D. On the Measurement of Polarization. *Econometrica* **62**, 819 (1994).
33. Almaatouq, A., Becker, J., Houghton, J. P., Paton, N., Watts, D. J., & Whiting, M. E. (2020). Empirica: a virtual lab for high-throughput macro-level experiments. *arXiv preprint arXiv:2006.11398*.
34. J. Hawes. Grace will lead us home: The Charleston Church Massacre and the Hard, Inspiring Journey to Forgiveness. *St. Martin’s Press*, *New York*. (2019).

**Acknowledgments:** Special thanks to Ray Reagans, Hazhir Rahmandad, Abdullah Almaatouq, Duncan Watts, Sinan Aral, David Rand, Carolyn Fu, Hagay Volvovsky, and the System Dynamics and Economic Sociology groups at MIT Sloan for their guidance.

1. along with commentary by Butts and others *(18-20)* [↑](#footnote-ref-2)
2. Throughout this paper, I use the term “beliefs” to refer to statements of fact, rather than values or ideals. [↑](#footnote-ref-3)
3. Avoiding the need to pre-specify compatibility between beliefs gives confidence that any systematic variation in adoption can be attributed to the diffusion process rather than to the assumed relationships. The importance of this is outlined in the supplement. [↑](#footnote-ref-4)
4. The results of the simulation are qualitatively similar with other thresholds or decision rules. [↑](#footnote-ref-5)
5. There are many complex measures of polarization in the literature *(14, 15, 17, 26-32 for a sample)*, which generally attempt to represent three basic intuitions. First, that individuals within the same ideological camp come to be more similar to one another. Secondly, that individuals in different ideological camps become more *dis*similar to one another. Lastly, that an individual’s position on one dimension of belief becomes informative of their position on other dimensions. As my purpose is not to identify camps and their members, but to suggest that one set of conditions is more generative of polarization than another, these measures add more complexity than value. Instead I report heuristic measures characterizing the above three intuitions. [↑](#footnote-ref-6)
6. This sample size was chosen primarily to satisfy budget constraints. [↑](#footnote-ref-7)
7. The preregistration can be found at <https://osf.io/239ns> [↑](#footnote-ref-8)
8. Indeed, this was one of the hopes of the early internet. Unfortunately, evidence suggests that the internet has not led to global concord. Present-day calls to break up online “filter bubbles” make a similar argument that changes to network structure will help reduce polarization, substituting only the type of change to be made. [↑](#footnote-ref-9)