AMERICAN SOCIOLOGICAL REVIEW

OFFICIAL JOURNAL OF THE AMERICAN SOCIOLOGICAL ASSOCIATION

ONLINE SUPPLEMENT
to article in
AMERICAN SOCIOLOGICAL REVIEW, 2018, VOL. 83

Beyond Social Contagion: Associative Diffusion and the Emergence of Cultural Variation

Amir Goldberg
Stanford University

Sarah K. Stein
Stanford University

PART A. PATH DEPENDENCE

What causes agents to gravitate in one direction over the other? In this supplement, we show that agents' preferences at equilibrium are path dependent and are not merely determined by agents' initial random preferences. The results plotted in Panel A of Figure 4 in the main text already point in that direction. The figure plots the final correlation between agents' preferences in a two-agent simulation, as a function of the initial correlation between their randomly generated preferences. As the plot shows, the final correlation is not determined by the initial correlation, except in extreme cases where the initial correlation is either strongly positive or strongly negative.

As illustration that initial correlation between agents' preferences does not determine the final correlation, we plot three randomly selected runs of the two-agent model in Panel A of Figure S1. The diagram plots the inter-agent preference correlation as a function of time (we plot only the first 200 iterations of the model for visualization purposes, as correlations tend to lock in beyond that point). The correlation patterns follow an erratic nonlinear path early on, often moving between negative and positive values. In one run, for example, the correlation increases beyond .5 before changing course and dropping toward –1. Eventually, all correlation patterns settle on a steady state once correlation nears 1 or –1.

What determines these changes in inter-agent preference correlations? Panel B in Figure S1 plots the magnitude of change in the inter-agent preference correlation between two subsequent model iterations as a function of the magnitude of change in agents' preferences (on a log scale). As we described in the main text, agents update their preferences in reaction to other agents' behaviors. The magnitude and direction of this update is random, drawn from a normal distribution. Agents retain this update only if it does not decrease constraint satisfaction. As the plot in Panel B demonstrates, shifts in preference correlations are almost entirely driven by the magnitude of changes in agents' preferences (r = .866). In other words, agents' stochastic preference updating behavior drives changes in their congruence with others.

Finally, to show that our results are not driven by initial inter-agent preference correlations, we ran a series of simulations of the two-agent model where agents' preferences are initialized to be zero for all preferences (namely, agents begin the simulation with neutral preferences, and all agents have the exact same preferences). Panel C in Figure S1 plots the proportion of negative and positive final inter-agent preference correlations (based on 1,000 simulations). Half the simulations tilt toward preference similarity, and half toward preference opposition. Thus, even when agents begin with identical preferences, the system still evolves stochastically. As the examples in Panel A illustrate, this stochasticity is not merely a function of the first several steps of the model.

Together, these three analyses show that our results are not simply determined by initial random preferences. Rather, the evolution of preferences and their inter-agent correlations are path dependent, driven predominantly by stochastic preference updates that are motivated by agents' desire to increase constraint satisfaction. Preferences determine outcomes only when their initial preference correlation is extremely strong.

PART B. NUMBER OF AGENT CLUSTERS AT EQUILIBRIUM

The multi-agent simulations consistently converge on a two-cluster equilibrium, where agents are partitioned into two opposing groups. As we explained in the main text, this happens because the interpersonal two-stage transmission process gradually imposes structural balance on agents' associative matrices. In this supplement, we provide more detail on how associative transmission leads to a two-cluster partition through structural balance, and we explore an alternative constraint function that produces a larger number of clusters.

In graph theory, a signed graph is defined as structurally balanced when it does not contain any cycles with one negative edge. For ease of exposition, let us assume that two practices represented in an agent's associative matrix can only be associated or dissociated. This is equivalent to a signed graph where nodes are connected by either unweighted positive or negative edges. Moreover, let us assume that agents have either unweighted positive or negative preferences.

To see how constraint satisfaction imposes structural balance on agents' associative matrices, consider a hypothetical agent's associative matrix, illustrated as a graph in Figure S2. Let a, b and c be three different practices. The two edges connecting the pairs $a - b$ and $b - c$ correspond to positive associations between these practices, respectively. The absence of an edge between a and c indicates that the two practices are dissociated. The plus signs indicate the agent has positive preferences for all three practices. This set of preferences is constraint satisfying for edges $a - b$ and $b - c$. However, it is not constraint satisfying for the absence of edge $a - c$, given that agents are forced to have different preferences for dissociated practices. To increase constraint, the agent can alter her preference for c to negative, such that she will have different preferences for the dissociated practices a and c; but that would be inconsistent with edge $b - c$, given that agents are constrained to have equal preferences for associated practices.

Overall, the only way constraint can be increased is by altering the agent's associative matrix and corresponding preferences to one of the four configurations illustrated in Figure S3 (Cartwright and Harary 1956). In our model, however, agents do not directly change their associative matrices to increase constraint. Rather, they only update associations in response to observed behaviors of other agents. Nevertheless, agents do update their preferences in response to these behaviors as a means to increase constraint satisfaction. Unlike the simple unweighted signed graphs in Figures S2 and S3, in our model agents' associative matrices are, effectively, non-negative weighted graphs. This allows for gradual changes in constraint satisfaction.

Ultimately, the mutually reinforcing dynamics of associative diffusion via interpersonal transmission lead to gradual increases in structural balance in agents' associative matrices. As Cartwright and Harary (1956) show, graphs that fully satisfy the four basic conditions illustrated in Figure S3 are perfectly balanced, leading to a partition into two clusters. That is why when there are no limitations on agents' ability to interact, associative diffusion results in cultural differentiation into two groups.

The four configurations illustrated in Figure S3 correspond to the four fundamental assumptions of balance theory. If the edges connecting nodes $a - b$ and $b - c$ are given, then the edge $a - c$ can only have the value specified in the diagram for the graph to be structurally balanced. As Davis (1967) shows, when assumption D (namely, if a and b are dissociated, and b and c are dissociated, then a and c must be associated; conventionally referred to as "the enemy of my enemy is my friend" condition) is relaxed, a balanced graph can have more than two clusters. To explore this possibility, we conducted an additional set of analyses in which we relax the constraint function to apply to associative matrix elements only when the agent's preferences for both practices are non-negative. To do so, we define:

$$CS_{relax}(V,R) = \frac{K}{K(K-1)} \sum_{i=1}^{K} \sum_{j=1}^{K} \delta(i,j) \left| R_{ij} - \Omega_{ij} \right| \quad \text{(S1)}$$

where:

$$\delta(i,j) = \begin{cases} 0, & \text{if } v_i < 0 \;\&\; v_j < 0 \\ 1, & \text{otherwise} \end{cases} \quad \text{(S2)}$$

This alternative to constraint satisfaction is analogous to a relaxation of condition D in balance theory, but it is not perfectly identical to it. It has the desired property of permitting agents to have equally negative preferences for dissociated practices, allowing these practices to be in different clusters. At the same time, it also leads to an undesired property where agents can have different levels of negative preferences toward strongly associated practices. Overall, when implemented with relaxed constraint, associative diffusion takes significantly longer to reach equilibrium with 30 agents (roughly at $t = 500{,}000$), resulting in 3.08 clusters, on average.

Although this result explains why our model converges on two clusters, and under what conditions more than two clusters can emerge, we are hesitant to reach any further conclusions, as it is not obvious to us that this implementation is consistent with how constraint operates cognitively. In particular, because this implementation relaxes the assumption of symmetric constraint satisfaction, it can lead agents to have dramatically different negative preferences for practices that they perceive to be strongly associated. This property seems inconsistent with how constraint operates. We leave further exploration of constraint relaxation and its effects on associative diffusion for future work.

PART C. ALTERNATIVE SPECIFICATIONS

In the main paper, we report the results of simulations modeling a variety of alternatives to associative diffusion using different network topologies. We show these alternatives cannot explain the emergence of cultural variation unless a segregated small-world network structure is assumed. In this supplement, we provide details on how we implement these different contagion mechanisms and network topologies.

Naive Contagion

All interpersonal transmission mechanisms assume contagion occurs when agent B's preference for practice i changes as a function of agent A's preference, as described in Equation 8. The basic interpersonal transmission mechanism, which we refer to as naive contagion, occurs when $V_{Bi}(t+1) = V_{Ai}(t)$. Most diffusion models in the literature assume that contagion is perfectly naive.

The diffusion dynamics generated by this simple contagion mechanism are uninteresting for our purposes, given that, in the absence of structural barriers to diffusion, they will always lead to cultural homogeneity. We introduce two additions to the transmission model. First, and drawing on existing literature (e.g., Dandekar, Goel, and Lee 2013; Friedkin and Johnsen 1990), we assume that a social susceptibility parameter α, ranging from 0 to 1, determines the extent to which agents are susceptible to influence. When $\alpha = 0$, agents are not affected by others' behaviors, whereas when $\alpha = 1$ they fully adapt their preferences to others'.

Second, we assume agents are unaware of others' private preferences. Rather, they observe others' behaviors and make inferences about their preferences. We define γ as the standard inference that an agent makes about another agent's private preference for a practice i when she observes that agent performing practice i. Together, we define naive contagion as follows:

$$V_{Bi}(t+1) = (1-\alpha)V_{Bi}(t) + \alpha\gamma \quad \text{(S3)}$$

When $\gamma = V_{Ai}(t)$ agents have full knowledge of their interlocutors' preferences. Such a model always leads to full preference convergence. When γ is fixed for agents, namely when agents always infer that another agent's preference is fixed, all preferences eventually (and unsurprisingly) converge toward γ.

To add stochasticity, we assume that γ is randomly and uniformly drawn from the range $[.1,1]$, that is, agents randomly infer other agents' preferences. The results plotted in Panel A of Figure 6 in the main text are based on a specification of $\alpha = .5$ and random γ. They are robust to different positive values of α.

In the specifications that follow, we always assume that γ ranges from .5 to 1 (i.e., agents infer a moderate to strong preference). The results reported in the main text are robust to this assumption and are reproduced when we assume agents have full access to others' preferences. Nevertheless, we believe a model assuming preference inference is more realistic than one in which agents have full access to others' preferences, especially when interaction is assumed to be superficial, as is the case in our model.

Biased Contagion

In the biased contagion condition, we model a process whereby agent B's preexisting preferences mediate the effects of A's behaviors. Following Dandekar and colleagues (2013), we implement biased contagion as a function of B's prior preference for i, weighted by a bias parameter $\beta > 0$. β defines the extent to which B's existing preferences mediate social transmission from A. As long as $\beta > 1$, bias is positive. Like Dandekar and colleagues (2013), we define biased contagion as a ratio between A's positive and negative effects on B's preference for i, weighted by β, as follows:

$$\breve{V}_{Bi}\,(t+1) = \frac{(1-\alpha)\hat{V}_{Bi}(t)+\alpha\hat{V}_{Bi}(t)^{\beta}\gamma}{(1-\alpha)+\alpha\hat{V}_{Bi}(t)^{\beta}\gamma+\alpha(1-\hat{V}_{Bi}(t))^{\beta}(1-\gamma)} \qquad \text{(S4)}$$

where α is again a social susceptibility parameter ranging from 0 to 1. In the results reported in the main text, we assume $\alpha = .5$, but these results are robust to different values of α as long as it is reasonably above 0 (roughly $\alpha > .1$) such that some social influence occurs.

Because the effect of bias is implemented as an exponentiation of B's existing preference, following Dandekar and colleagues (2013) we transform this preference to a 0 to 1 range using the logistic function, such that $\hat{V}_{Bi}\,(t) = \frac{1}{1+e^{-V_{Bi}(t)}}$. We then transform $\breve{V}_{Bi}\,(t+1)$ back into an infinite range using the logit function, $V_{Bi}(t+1) = \log\frac{\breve{V}_{Bi}(t+1)}{1-\breve{V}_{Bi}(t+1)}$ so that it takes negative and positive values, in compliance with our model's assumption about the range of preference values. This functional form has the desired behavior, such that values above 0 for B's prior preference lead to a growing positive effect on that preference, and those below 0 to a growing negative effect. The results we report in Figure 6 in the main text are robust to different values of γ and β.

Conformist Contagion

In the conformist contagion condition, we model a process where B's preference for i is mediated by B's taste for popularity and her perception of practice i's rarity. To do so, we define two additional parameters. First, we define ω_B as $B's$ taste for popularity, ranging from 0 to 1. We assign agents with a random taste for popularity, drawn from the inverse of a log normal distribution with a mean of $\log .15$ and standard deviation of $\log 2$. This ensures that the majority of agents are conformist (with half of agents having a taste for popularity at or greater than .85), and a minority are nonconformist. Second, we define B's perception of practice i's rarity as a function of how frequently she has observed other agents performing that practice. For each agent, we therefore define a *K*-sized vector O_B that is initialized to 0, and where the value of cell i increases by 1 whenever B observes that practice enacted by others. Values in O decay as a function of a decay parameter λ. We can now define $\psi_{Bi} = 1 - \frac{O_{Bi}}{\max(O_B)}$ as B's perception of practice i's rarity.

Building on Flache and Macy (2011), we define conformist contagion as follows:

$$V_{Bi}(t+1) = V_{Bi}(t) + (2 \cdot |\omega_B - \psi_{Bi}| - 1)\gamma \qquad \text{(S5)}$$

This mechanism of contagion ensures that B's preference changes as a function of the distance between her taste for popularity and her perception of the practice's rarity, $|\omega_B - \psi_{Bi}|$. As that distance nears 1, that is, as the congruence between the practice's perceived rarity and the agent's taste for popularity grows, B increases her preference for practice i. As the distance nears 0, B decreases her preference for i. The results reported in Figure 6 in the main text are robust to different values for γ, as long as it is positive.

We also consider an alternative method of assigning tastes for popularity, where we dichotomously divide the population into conformists and nonconformists. We define the tastes for popularity for these two conditions as $\omega = .95$ (conformist) and $\omega = .05$ (nonconformist), and we randomly assign 75 percent and 25 percent to each of these two conditions, respectively. Unlike the log-normal method, which generates a skewed distribution of taste for popularity, this dichotomous method generates a bimodal distribution. The results reported in Panel A of Figure 6 in the main text are robust to this specification, suggesting that even when there is a clear division into conformists and

nonconformists (such as in the case of early adopters who are distinctively and qualitatively different from mainstream audiences), a division into different cultural clusters does not emerge.

In Panels B and C of Figure 6 in the main text, we explore the effects of extending the baseline associative diffusion model to account for variation in conformity (as specified in Equation 9). Panel B reports results from this model, where taste for popularity is generated using the log-normal method as detailed earlier. In Panel C, we run multiple simulations of the extended associative diffusion model where we vary the overall prevalence of conformity. To do so, we use the dichotomous method for generating taste for popularity, and we vary the proportion of conformists in each simulated run. For example, when the proportion of conformists is .6, 60 percent of agents have a high taste for popularity at $\omega = .95$, and the remaining 40 percent have a low taste for popularity at $\omega = .05$. As the plot shows, cultural differentiation emerges as long as the proportion of conformists is greater than roughly .15.

Homophily

In the homophilous contagion condition, we model a process where B's change in preference for i is mediated by B's perceived homophily with A. Consistent with our assumption of superficial interaction, we assume B only has partial information about A's other preferences. Specifically, B observes A perform only one additional practice, j. We define B's homophily with A as the perceived similarity between their preferences for j.

Diffusion models that take into account the effects of cultural similarity on adoption normally calculate similarity in Euclidean space (e.g., Baldassarri and Bearman 2007). We build our implementation of homophily on DellaPosta, Shi, and Macy (2015). DellaPosta and colleagues' model calculates the similarity between agents B and A as the difference between two distances: the Euclidean distance between the agents and the expected distance between two random agents drawn from the population. The probability of social influence, or the likelihood that B will adopt A's preference, is proportional to the magnitude of that difference. In other words, the stronger the agents' similarity or dissimilarity (relative to what would be expected at random), the greater the likelihood of social influence. When the agents are dissimilar (i.e. their distance is greater than expected) adoption is negative. But because negative influence is rarer than positive influence, it randomly occurs in only 10 percent of cases.

We adapt this model to our setting, where preferences range from negative to positive values and where preferences are only partially observable. Although, unlike in DellaPosta and colleagues' model, our agents do not observe others' full set of behaviors, we assume they are aware of the private preferences for the behaviors they observe. We therefore define the social influence of A on B as the inverse of the absolute distance between their preferences for practice j:

$$W_{Bj,Aj}(t) = 1 - ||V_{Bj}(t)| - |V_{Aj}(t)|| \qquad \text{(S6)}$$

Because preferences are initially drawn from the range –1 to 1, $W_{Bj,Aj}$ ranges from 0 to 1. Equation S6 ensures that preferences remain within the –1 to 1 range. We define homophilous contagion as follows:

$$V_{Bi}(t+1) = (1 - W_{Bj,Aj})V_{Bi}(t) + W_{Bj,Aj}V_{Ai}(t) \qquad \text{(S7)}$$

When A's and B's preferences for j are differently signed, that is, one has a positive and the other a negative preference, we define homophilous contagion as follows:

$$V_{Bi}(t+1) = (1 - W_{Bj,Aj})V_{Bi}(t) - W_{Bj,Aj}V_{Ai}(t) \qquad \text{(S8)}$$

We allow such influence to occur in only 10 percent of cases, as per DellaPosta and colleagues (2015).

This implementation ensures that, like in DellaPosta and colleagues' model, equally similar and dissimilar agents have influence of the same magnitude, but in opposite directions. Panel A of Figure 7 in the main text reports results using this model.

To ensure the results reported in Figure 7 in the main text are not driven by specific assumptions, we also examine an alternative specification for the homophilous contagion process, which builds on and extends Flache and Macy

(2011). In this specification, we assume agents do not have access to interlocutors' private preferences. We define alternative homophilous contagion as follows:

$$V_{Bi}(t+1) = V_{Bi}(t) + [2(1 - \Delta_{Bj}(t)) - 1]\Delta_{Bi}(t) \quad \text{(S9)}$$

where $\Delta_{Bi}(t) = \gamma - \widehat{V}_{Bi}(t)$ is the distance between γ, which is what B infers as A's preference for an enacted practice, and B's own preference for that practice (transformed to the 0 to 1 range as explained earlier, to comply with Flache and Macy's [2011] model). This functional form ensures that as B's preference for j grows closer to her inference about A's preference, she updates her preference for i to be increasingly identical to her inference about A's preference. When her preference for j is significantly lower than γ, this update rule means she decreases her preference for i. Results using this specification replicate the results reported in Figure 7.

Network Topologies

We also explore associative diffusion in three different network topologies: a fully connected network, a scale-free network, and a small-world network. We implement these different topologies as directed graphs (with no self-edges). In each simulation round, one observer agent, B, is randomly selected with uniform probability. The actor agent, A, is selected with uniform probabillity from the subset of agents to whom B has an outgoing edge. In this subsection, we explain how the three network topologies are generated.

The fully connected topology is a graph in which all potential edges are realized (with the exception of self-edges connecting a node to itself). Such a network implies that all edges have a uniform probability (equal to $\frac{1}{N(N-1)}$) of being selected. This is equivalent to selecting an observer and actor with uniform random probability.

A scale-free network is one in which node indegree follows a power law distribution such that the probability of nodes with k incoming edges, $P(k) \sim k^{-\alpha}$. We generate networks where α ranges from two to three and where each node has an outdegree of six. To generate such a network, we randomly assign all nodes with a popularity score that follows a power law distribution. We then iterate over all nodes, and we assign them with six random outgoing edges to other nodes with a probability proportional to these nodes' popularity. Such a process generates a network wherein each agent can observe only six other agents, but agents vary significantly in how many agents can observe them.

A small-world network topology is one in which nodes are segregated into clusters. Following Watts (1999), we implement a connected caveman topology with five clusters. To do so, we randomly divide the network into five equally sized and fully connected cliques. We then randomly rewire 10 percent of the edges in the network, by randomly selecting two edges and swapping their destination nodes. This generates edges that bridge between different cliques. The procedure we follow is similar to the procedure in DellaPosta and colleagues (2015).

An important feature of the network generation processes we implement is that both the scale-free and small-world networks have the same overall number of edges and the same node outdegrees. That is, in both types of networks, each agent can observe the same number of other agents. These networks differ, however, in how these edges are distributed. In the scale-free topology, a small number of agents account for the majority of indegrees. In the small-world topology, indegrees are equally distributed.

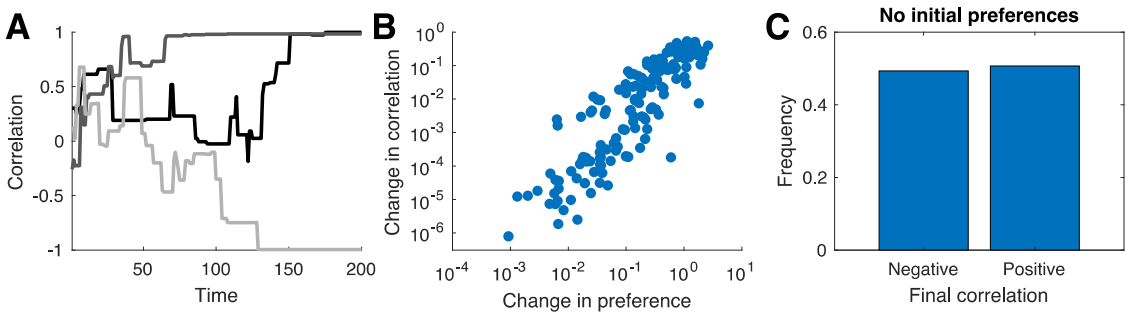

Figure S1. Path Dependence in the Two-Agent Model

Note: (A) Examples of the evolution of inter-agent preference correlation from three random simulation runs. (B) Changes in inter-agent preference correlation as a function of stochastic preference updates. (C) The proportion of negative and positive final inter-agent preference correlations when agents are initialized to have uniform 0 preferences.

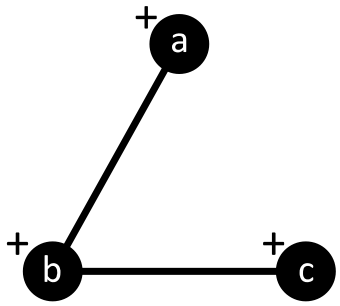

Figure S2. Illustration of Structural Imbalance in a Triad Contained in an Agent's Associative Matrix

Note: The three nodes correspond to practices, which are either connected by an edge (associated) or not (dissociated). Plus signs represent the agent's positive preference toward these practices.

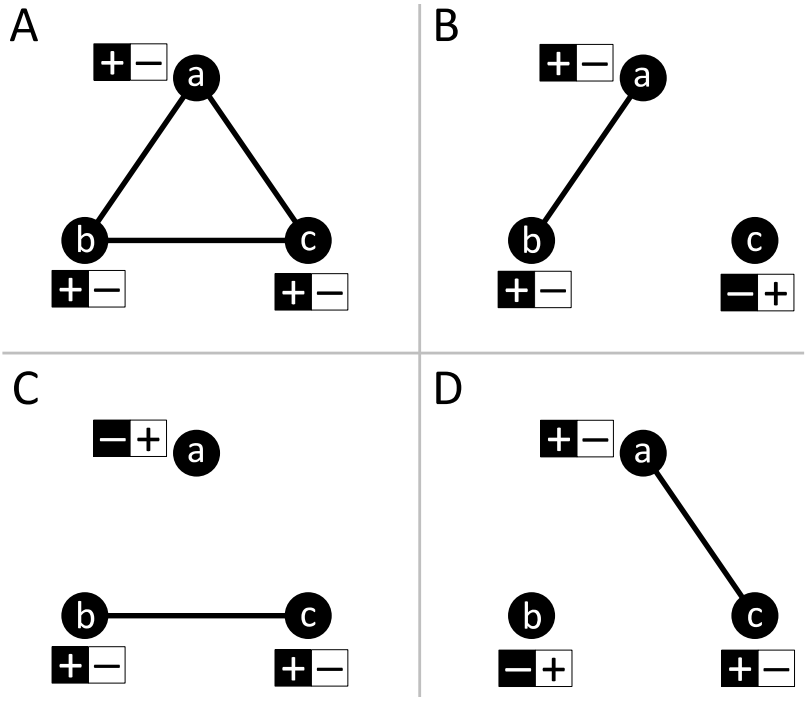

Figure S3. Four Configurations of Balanced Triads in an Agent's Associative Matrix

Note: Nodes correspond to practices, which are either connected by an edge (associated) or not (dissociated). The boxed plus and minus signs correspond to the preferences the agent must have in order to reach balance. For each configuration, preferences must adhere to either the black or white boxes to satisfy constraint. For example, in Panel A, the agents must have either positive or negative preferences for all practices.

References

Baldassarri, Delia, and Peter Bearman. 2007. "Dynamics of Political Polarization." *American Sociological Review* 72(5):784–811. https://doi.org/10.1177/000312240707200507.

Cartwright, Dorwin, and Frank Harary. 1956. "Structural Balance: A Generalization of Heider's Theory." *Psychological Review* 63(5):277–93. https://doi.org/10.1037/h0046049.

Dandekar, Pranav, Ashish Goel, and David T. Lee. 2013. "Biased Assimilation, Homophily, and the Dynamics of Polarization." *Proceedings of the National Academy of Sciences* 110(15):5791–6. https://doi.org/10.1073/pnas.1217220110.

Davis, James A. 1967. "Clustering and Structural Balance in Graphs." *Human Relations* 20(2):181–87. https://doi.org/10.1177/001872676702000206.

DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why Do Liberals Drink Lattes?" *American Journal of Sociology* 120(5):1473–1511. https://doi.org/10.1086/681254.

Flache, Andreas, and Michael W. Macy. 2011. "Small Worlds and Cultural Polarization." *Journal of Mathematical Sociology* 35(1–3): 146–76. https://doi.org/10.1080/0022250X.2010.532261.

Friedkin, Noah E., and Eugene C. Johnsen. 1990. "Social Influence and Opinions." *Journal of Mathematical Sociology* 15(3–4):193–206. https://doi.org/10.1080/0022250X.1990.9990069.

Watts, Duncan J. 1999. "Networks, Dynamics, and the Small-World Phenomenon." *American Journal of Sociology* 105(2):493–527. https://doi.org/10.1086/210318.