3-2015

# The Social Origins of Networks and Diffusion

Damon Centola
*University of Pennsylvania*, dcentola@asc.upenn.edu



## Recommended Citation

Centola, D. (2015). The Social Origins of Networks and Diffusion. *American Journal of Sociology, 120* (5), 1295-1338. https://doi.org/10.1086/681275

## The Social Origins of Networks and Diffusion


### Abstract

Recent research on social contagion has demonstrated significant effects of network topology on the dynamics of diffusion. However, network topologies are not given a priori. Rather, they are patterns of relations that emerge from individual and structural features of society, such as population composition, group heterogeneity, homophily, and social consolidation. Following Blau and Schwartz, the author develops a model of social network formation that explores how social and structural constraints on tie formation generate emergent social topologies and then explores the effectiveness of these social networks for the dynamics of social diffusion. Results show that, at one extreme, high levels of consolidation can create highly balkanized communities with poor integration of shared norms and practices. As suggested by Blau and Schwartz, reducing consolidation creates more crosscutting circles and significantly improves the dynamics of social diffusion across the population. However, the author finds that further reducing consolidation creates highly intersecting social networks that fail to support the widespread diffusion of norms and practices, indicating that successful social diffusion can depend on moderate to high levels of structural consolidation.



### Disciplines

Communication | Social and Behavioral Sciences

# The Social Origins of Networks and Diffusion[1]

Damon Centola
*University of Pennsylvania*

Recent research on social contagion has demonstrated significant effects of network topology on the dynamics of diffusion. However, network topologies are not given a priori. Rather, they are patterns of relations that emerge from individual and structural features of society, such as population composition, group heterogeneity, homophily, and social consolidation. Following Blau and Schwartz, the author develops a model of social network formation that explores how social and structural constraints on tie formation generate emergent social topologies and then explores the effectiveness of these social networks for the dynamics of social diffusion. Results show that, at one extreme, high levels of consolidation can create highly balkanized communities with poor integration of shared norms and practices. As suggested by Blau and Schwartz, reducing consolidation creates more crosscutting circles and significantly improves the dynamics of social diffusion across the population. However, the author finds that further reducing consolidation creates highly intersecting social networks that fail to support the widespread diffusion of norms and practices, indicating that successful social diffusion can depend on moderate to high levels of structural consolidation.

Recent theoretical and empirical studies on the spread of social norms have identified the importance of network topology—the large-scale pattern of ties within a population—for determining both the rate and extent of diffusion processes (Watts and Strogatz 1998; Newman and Watts 1999; New-

[1]The author gratefully acknowledges suggestions and guidance from Duncan Watts during the early stages of this project and extensive comments from Paul DiMaggio, Arnout van de Rijt, Peter Dodds, Michael Macy, Ezra Zuckerman, Ray Reagans, and Ronald Burt. This research was supported in part by the James S. McDonnell Foun-


0002-9602/2015/12005-0001$10.00

man, Barabasi, and Watts 2006; Centola and Macy 2007; Centola 2010). While much of the literature on networks treats the topological features of social structure as given quantities, networks do not emerge *ex nihilo*, but are endogenously formed patterns of relationships that are created by individual and institutional forces that constrain and direct everyday interactions (Blau 1977; McPherson and Smith-Lovin 1987; McPherson, Popielarz, and Drobnic 1992; Popielarz and McPherson 1995; Watts, Dodds, and Newman 2002; Kossinets and Watts 2009). For instance, at the individual level, choice homophily—people's preference to form social connections with others who are like themselves—is a powerful force that controls the formation of social ties (Verbrugge 1977; Blau and Schwartz 1984; Shrum, Cheek, and Hunter 1988; McPherson, Smith-Lovin, and Cook 2001; Reagans 2005; Currarini, Jackson, and Pin 2010; Centola and van de Rijt 2015). At the level of social structure, organizational and institutional contexts form the basis for social interaction (Blau 1977; McPherson et al. 1992; McPherson 2004), and the distribution and consolidation of characteristics across a population determine the frequency with which people with diverse characteristics come in contact with one another (Blau 1977; Blau and Schwartz 1984; McPherson et al. 1992). Together, these forces help to shape the collective pattern of network ties that emerges within a society.

These individual and institutional forces can thereby indirectly affect the potential for social integration. Classic work by Blau and Schwartz (1984) on the relationship between social institutions and emergent network structures found that "consolidated" societies—where one's friends, colleagues, and neighbors are all the same people—limit social integration across the society by preventing the formation of "cross-cutting social ties" between diverse social groups. I take Blau and Schwartz's structural theory one step further by investigating social integration as a diffusion process—in particular, the spread of shared cultural practices and social norms throughout a population (Boyd and Richerson 1985; Castro and Toro 2004; Centola and Macy 2007). The central goal of this article is to explore how changes in a society's population structure affect these forms of social diffusion.

I proceed by first investigating how the "primitive" structural parameters of homophily and consolidation (Skvoretz 1983; Blau and Schwartz 1984) control the formation of social networks with distinct topological properties—such as clustering, path length, and bridge width. I then explore the effectiveness of these emergent networks for promoting the spread of

dation through a Studying Complex Systems grant to the author. Direct correspondence to Damon Centola, Annenberg School of Communications, University of Pennsylvania, 3620 Walnut Street, Philadelphia, Pennsylvania 19104. E-mail: dcentola@asc.upenn.edu

social norms. I find that social diffusion depends on consolidation in surprisingly complex ways.

In contrast to Blau and Schwartz's findings that social consolidation has a monotonically dampening effect on social integration, my results show an inverted U-shaped effect of consolidation on social diffusion. As Blau and Schwartz (1984) observed, I find that consolidation can be too strong, which limits social integration by fragmenting a society into isolated groups. However, I also find that consolidation can be too weak, preventing the formation of wide network bridges that coordinate and reinforce the adoption of new ideas, norms, and behaviors across a population. In between these extremes, I find a region of surprisingly high consolidation where social integration on shared norms easily succeeds. The success of these networks derives from the community structures that emerge, in which mesolevel patterns of overlapping groups, connected through wide bridges, establish the necessary social fabric to support the spread of shared norms and practices throughout a population.

My conclusions inform demographic and organizational thinking about how social networks affect cultural integration and suggest new methodological connections between traditional population-based research and recent computational work on network dynamics.

## CONSOLIDATION AND HOMOPHILY

Blau (1977; extending Simmel [(1922) 1955]) was among the first scholars to develop a formal theory of how features of social structure control the formation of social ties. Building on this work, Blau and Schwartz (1984) explicitly connect their research on homophily in the formation of social ties to the deeper "topological" question of how systematic large-scale patterns of relationships emerge from the interaction of organizational constraints and individual preferences. Arguing that homophily alone is insufficient to explain network structure, Blau and Schwartz emphasize the central role of consolidation in determining the collective properties of emergent social networks (see also Skvoretz 1983). As Blau defines it, consolidation is the correlation between traits in a population. For instance, if race and income are highly correlated, then knowing someone's race implies that you can accurately predict her income. The more consolidated a society is, the more variables are correlated with one another. This implies that in a highly consolidated society, knowing a single property of an individual (e.g., race) would be sufficient to predict most of her other characteristics (e.g., income, education, health status, etc.).

Blau and Schwartz's (1984) study of social structure concludes that the norms and institutions that promote social consolidation create patterns of social ties that reduce social integration. For Blau and Schwartz, integration is defined in terms of social connections between diverse individuals in a social network; that is, it is a structural feature of the network configuration of a society (p. 15). Consolidation reduces integration because the networks that emerge from consolidated societies are highly segregated, with very few crosscutting ties. The goal of the present study is to expand on this work by asking a related question, namely, How does social consolidation affect the diffusion of shared norms and customs across large populations?

I investigate social integration as a diffusion process, that is, as the spread of social "contagions," such as shared practices and beliefs. For "simple" social contagions, such as rumors, information about jobs, breaking news, and even the influenza virus, transmission requires only a single contact. Once someone tells me the score of the game or a piece of breaking news, I can easily repeat it to others. However, for the diffusion of "complex" social contagions, such as norm compliance, cultural practices, and costly collective action, people frequently require social reinforcement from more than one other adopter in order to be convinced to adopt the behavior themselves (Boorman 1974; Friedkin 1998; Centola and Macy 2007). Building on this tradition of social research (Granovetter 1978; Friedkin 2001), I study the dynamics of social integration as the spread of complex social contagions across the social landscape. Through a series of computational investigations, I analyze the full spectrum of social networks that emerge from the interaction of homophily and consolidation. I find that consolidation has a surprisingly curvilinear effect on the dynamics of social diffusion.

As Blau and Schwartz (1984) suggest, very high levels of consolidation do indeed create social boundaries, which prevent behaviors from spreading across a population. As social consolidation is reduced, the formation of crosscutting social circles promotes the dynamics of social integration through the diffusion of collective behaviors. However, if social consolidation is reduced too much, the networks that emerge can have a surprisingly inhibitory effect on the spread of collective behaviors. While too much consolidation creates a balkanized world of cultural segregation, highly intersecting networks prevent groups from forming coherent collective practices. I find that moderate levels of social consolidation within a society can actually strengthen the collective social network, creating overlapping group structures that support the widespread integration of shared norms and beliefs across a population.

## HOW CONSOLIDATION AND HOMOPHILY AFFECT SOCIAL NETWORKS

Blau and Schwartz's (1984) study of crosscutting social circles begins with the idea that there are "dimensions" of social life, such as work, neighborhood, volunteer organization, and so on. Each dimension of social life defines a context for social interaction, and each context provides an opportunity to form social ties.[2]

In a relatively undifferentiated society, where there is not much diversity among jobs or neighborhoods, people's interactions are not highly constrained by their social positions (Blau and Schwartz 1984). Ties form indiscriminately between people, who all mix and mingle in the same social "soup." However, as the division of labor in society increases and organic solidarity replaces mechanical solidarity (Durkheim [1893] 1997), the greater diversity in the number of social positions in the workplace comes with attendant changes in the number of neighborhoods, organizations, and clubs to which people can belong. In a large, diverse society, each dimension of social life can have much greater heterogeneity among social positions.

Social positions that are similar to one another are "proximate" within a dimension of social life, which implies a greater likelihood of social contact between their members than with members of more "distant" positions; for example, the police officer is "closer" to the fireman than he is to the engineer and thus more likely to develop a work tie with the fireman. This also suggests that the more differentiation there is within a dimension of social life, the greater the social distance can be between people. For this reason, Blau and Schwartz (1984) emphasize that greater heterogeneity in the number of social positions in a society has important implications for how social networks form. Because the formation of social ties is constrained by opportunities for interaction, greater social distance between people implies that they will have few or no opportunities to form ties with one another. Consequently, underlying changes in the level of heterogeneity in a population can have profound effects on its emergent patterns of social relations.

Blau and Schwartz's (1984) study focuses on three key parameters:

1. heterogeneity—the number of social positions in each dimension of social life,
2. inequality—skewness of the distribution of status and resources across social positions,
3. consolidation/intersection—the degree to which people's social position in one dimension of social life correlates with their position in other dimensions.

Strikingly, Blau and Schwartz's core finding is that the most salient structural parameter for understanding the emergent pattern of network ties is social consolidation. As Blau and Schwartz put it, "the degree to which

[2] Other related work has fruitfully referred to social dimensions as "contexts" and "foci" for social interaction (Feld 1981; McPherson and Smith-Lovin 1987; McPherson 2004). For consistency with Blau and Schwartz, I use the term "dimension" in the present article.

social differences intersect [i.e., the degree to which they are not consolidated] is of prime significance for intergroup relations and a community's integration. Intersection is the central concept of the theory under consideration" (p. 12).

To understand Blau and Schwartz's emphasis on consolidation (and its complementary concept of intersection), it is useful to consider an example. Consider a society in which there are well over a hundred different categories of possible employment, and the distribution of memberships across those different social positions is relatively equal. Thus, each social position has approximately the same number of members: for example, there are an equal number of bakers, bankers, bus drivers, and so forth in the population. What does this tell us about the likelihood that a baker is the friend of a banker? And what does that imply about the social and cultural norms shared by bakers and bankers? What Blau and Schwartz (1984) find is that in order to answer these questions, it is necessary to look at the other (nonwork) dimensions of social life and how those dimensions are related to one another.

While some social ties are made at work, through interactions with colleagues, other ties are made at home, interacting with neighbors, and yet other ties are made at the lodge, or country club, interacting with fellow members of civic and voluntary organizations. Thus, to know how likely it is that a baker is friends with a banker, it is necessary to know whether the bakers tend to live in the same neighborhood as bankers, whether they tend to belong to the same lodges or join the same sports leagues.

Looking across these different dimensions of social life, one key structural parameter becomes salient: How likely is it that a person's position in one dimension will correlate with her position in other dimensions? If I know your job, can I reasonably predict where you live, what civic organization you belong to, and where you dine? The more correlated these dimensions of social life are with one another, the more consolidated a society is. By contrast, if knowing a person's location in one of these dimensions tells us very little about her location in the other dimensions, then a society has low consolidation (i.e., it is highly "intersecting"). Intuitively, the lower the level of consolidation, the greater the likelihood that people with highly distant positions in one dimension (e.g., the banker and the bus driver) will actually share positions in some other dimension (e.g., frequent the same sports bar) and therefore have some opportunity to interact with one another socially. Blau and Schwartz show that lower levels of consolidation produce crosscutting social circles—social ties between members of diverse social groups, which facilitate widespread social integration.

The intuitive clarity of this finding notwithstanding, it is not correct unless we make one important assumption. That assumption is that social interactions are constrained by homophily. The principle of homophily is simply that "likes" attract: We tend to interact with people who share social characteristics with us (Simmel 1950; Blau and Schwartz 1984; McPherson and Smith-Lovin 1987; McPherson et al. 2001). This assumption is essential for Blau and Schwartz's insight into the effects of social consolidation because even if all the bankers live in the same neighborhood, join the same clubs, and shop at the same stores, if social ties are not dependent on these contextual constraints, social distance will not matter for social relations. If people simply make ties at random without needing to have some common shared context, then structural constraints on social life do not affect who is likely to befriend whom. Without homophily, social consolidation does not matter.

But, of course, social ties are not made at random. Our embeddedness in social contexts reflects our interests, appetites, and ambitions, which constrain the social contacts we have, the friends we make, and (unbeknownst to us) the shape of the social networks in which we live and act (Simmel 1955; Blau and Schwartz 1984; McPherson et al. 1992, 2001; Popielarz and McPherson 1995; Centola et al. 2007). Somewhat more formally, the assumption of homophilous interaction can be stated as follows:

1. "Social associations are more prevalent between persons in proximate than those in distant social positions" (Blau and Schwartz 1984, p. 27).
2. "Rates of social association depend on opportunities for social contact" (Blau and Schwartz 1984, p. 29).

These postulates suggest that people's social ties are constrained by their location in "social space" (McPherson et al. 1992; Popielarz and McPherson 1995; McPherson 2004; Kossinets and Watts 2009). People who are closer within a dimension of social life will be more likely to interact with each other than those who are far apart. Blau and Schwartz's (1984) insight into the role of consolidation suggests that we need to understand not only how relations are constrained homophilously within each dimension but also how the dimensions themselves are correlated with one another.

The combination of homophily—the formation of social ties to people who are proximate within a dimension of social life—and consolidation—the correlation of social positions across dimensions—provides an analytic framework for modeling how social structures give rise to patterns of social relations. The strategy of what follows is to show how the parameters of homophily and consolidation interact to produce emergent social networks with distinctive topological properties. However, I am interested not only in how these social forces affect the emergence of different kinds of social networks but in how the resulting networks in turn affect the diffusion of shared beliefs and behaviors across the population.

In particular, a long tradition in the social diffusion literature has argued that "weak ties" between remote areas of a social network can promote the spread of new ideas and behaviors (Granovetter 1973; Rogers 1995; Watts and Strogatz 1998; Centola and Macy 2007). Granovetter's (1973) study of the strength of weak ties and, more recently, Watts and Strogatz's (1998) model of "small-world" networks both support the idea that reduced social clustering and increased weak ties across a population will promote social diffusion.[3] My strategy in what follows is to connect these ideas with Blau and Schwartz's (1984) argument that the emergence of crosscutting social ties that link across a diverse population is directly controlled by the level of consolidation within a society. The more consolidated a population is, the fewer crosscutting ties there are. Complete consolidation produces disconnected, homogeneous groups of people—for example, the rigidly stratified "company towns" that populated the American West in the 19th and early 20th centuries—while zero consolidation produces highly intersecting social networks in which people's characteristics provide very little constraints on with whom they interact—for example, the ideal of the highly crosscutting, rapidly diffusing social networks that constitute social life in the modern metropolis (Simmel 1950; Blau and Schwartz 1984). The implication that I hope to draw out from these connections is that potentially minor changes to social institutions that reduce or increase the level of social consolidation within a society can be unintentionally amplified through the vehicle of social networks into significant consequences for a population's collective capacity for social diffusion.

## A MODEL OF HOW SOCIAL NETWORKS EMERGE FROM THE INTERACTION OF HOMOPHILY AND CONSOLIDATION

My approach to studying network formation builds on the formal architecture developed by Watts et al. (2002), in which individuals not only have social ties but also have identities, which define their "social proximity" to other members of the population. I extend this formal architecture in two ways. First, I introduce social consolidation into the network construction algorithm, which allows me to study how the interaction between homophily and consolidation affects the topological properties of the emergent social networks. Second, I combine this network generation model

[3] Related work in computer science (Kleinberg 2000), epidemiology (Keeling 1999; Liljeros et al. 2001), and physics (Newman 2000) has echoed these findings, showing how networks with randomly placed long-distance ties can be more effective for the dynamics of social diffusion processes. Further, recent online studies have begun to explore how these structural properties affect communication (Albert, Jeong, and Barabasi 1999; Dodds, Muhamad, and Watts 2003) and influence dynamics across virtual networks (Backstrom et al. 2006; Centola 2010, 2011).

with Centola and Macy's (2007) diffusion model of complex contagions to study how the generative mechanisms underlying complex social topologies affect the dynamics of social diffusion.

Following Blau and Schwartz (1984) and Watts et al. (2002), my approach is based on the idea that there are different "dimensions" of social life, and each individual has a "social position" within each of those dimensions. For instance, "work" is a dimension of social life, and each individual has a position within this dimension, which corresponds to the job that he or she has. Unlike pure categorical variables for which each value can be treated as equidistant from every other position, the different positions in the employment dimension have an ordinal relationship to one another. At one end, skilled jobs require years of training and command high salaries; at the other end, low-paying jobs attract unskilled labor. Another example is the "neighborhood" dimension, which corresponds to the physical location where each individual lives. Like employment, this dimension of social life is stratified such that one end of the spectrum corresponds to inexpensive, less attractive neighborhoods, while the other end corresponds to "high-end" neighborhoods.

There are many different dimensions of social life that can be considered in addition to a person's job and neighborhood, such as club or lodge memberships, educational affiliations, entertainment venues, and so forth. In the model that I propose, each individual has a single position within each dimension, and each dimension provides an opportunity for social interaction. For example, in the employment dimension, people interact with, influence, and befriend their colleagues. Thus, a certain fraction of each individual's social ties will be contacts made at work. This is also true for the other dimensions since each dimension is an opportunity for individuals to make social ties. An individual's complete "ego network" (the total number of social ties that an individual has) is the sum of her ties across all the dimensions of social life.

Intuitively, this means that for each person, some of her social ties will be colleagues from work, some of them will be neighbors, some will be friends from the lodge, some will be people from school, and some will be contacts from parties, sporting events, concerts, and so forth. Of course, sometimes people may wind up seeing the same individuals in more than one dimension of their lives: people from my school, or my major, may also be members of the same professional societies as I am. The extent to which people show up in the same social positions across different dimensions of social life is determined by the level of social consolidation.

The greater the level of consolidation, the narrower the subset of the population to whom each person is exposed. When there is perfect consolidation, the same people show up in the same social positions across every dimension of life: A person's colleagues are also her neighbors, who are also

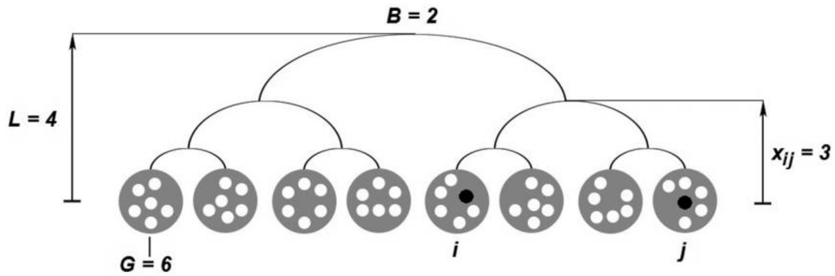


FIG. 1.—Social identities in a dimension of social life. Each social group is populated by $G$ members of the population. The distance between two individuals $i$ and $j$ is determined by how many steps up the branching tree it is to reach a common "ancestor." Members of the same social group are at a distance of 1, which is the minimum distance in the social space.

her fellow club members, who are also her classmates from school, and so forth. In such a world, we expect the emergent social networks to be highly segregated since each person's set of possible acquaintances is nonoverlapping with those of members of different groups.

### Social Position and Social Distance

To formalize Blau and Schwartz's (1984) concepts of social dimensions and social positions, I begin by considering the simple social structure shown in figure 1. The schema in figure 1 shows a representation of a single dimension of social life that has eight social positions.[4] Each of the ovals corresponds to a single social position (ranked from 1 through 8), and the circles within the ovals represent the individuals located in those positions. Each position has six members. For example, for the employment dimension, figure 1 indicates that there are eight categories of work, each of which is populated by six individuals. For example, figure 1 indicates that there are six senior executives (position 1), six junior executives (position 2), six middle managers (position 3), and six lower managers (position 4). These four positions on the left side of figure 1 constitute "management." Similarly, there are also four positions on the right side representing "labor," each with six members, ranging in rank from foreman (group 5) down to janitor (group 8).

Figure 2 expands the picture in figure 1 to include additional dimensions of social life. It shows four identically structured dimensions, each containing eight social positions. As above, each position has six members. Since each individual has one position in each dimension, each of the 48 individuals shown in figure 2 has four distinct social positions that define her

[4] The complete formal model for the schemas shown in figs. 1 and 2 is provided in the appendix.

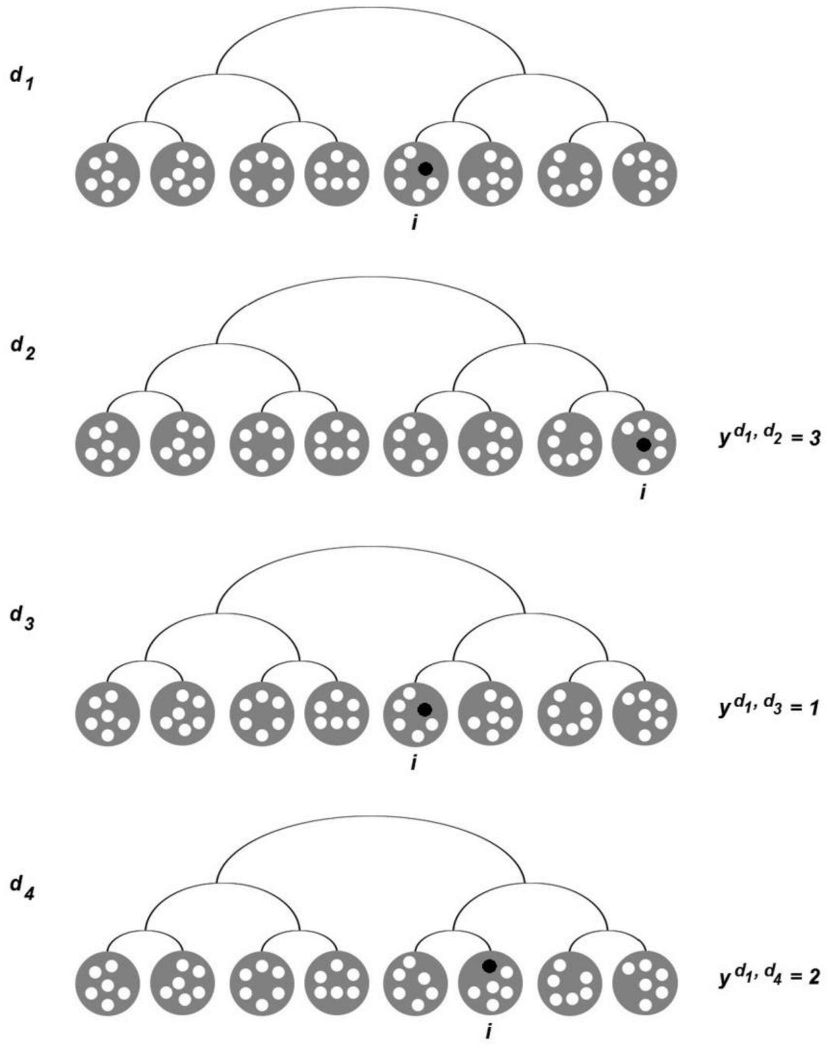


FIG. 2.—Multiple dimensions of social life. Every member of the population has a location in each dimension of social life. The distance between an individual $i$'s position across dimensions is evaluated using the same metric that evaluates the distance between two individuals, $i$ and $j$, within a single dimension: The distance between $i$'s social positions across different dimensions is the number of steps up the branching tree to find a common ancestor. When $i$ is located in the same group across social dimensions, the distance between positions is 1, which is the minimum distance.

social identity (e.g., a job, a neighborhood, a club membership, and a school affiliation).

Figure 2 shows that in dimension 1 ($d_1$), individual $i$ is in position 5. In $d_2$ she is in position 8; in $d_3$, position 5, and in $d_4$, position 6. If $i$ were in the same position across every dimension, the population would be perfectly consolidated since knowledge of $i$'s position in one dimension would provide full information as to her position in the other dimensions.

Social distance between individuals within a dimension of social life is measured using the "ultrametric" scale employed by Watts et al. (2002).[5] As shown in figure 1, the social distance between two individuals $i$ and $j$, denoted as $x_{ij}$, within a given dimension is measured by counting the steps between their respective social positions in terms of the closest fork in the social tree where they share a common branch. For example, if two individuals $j$ and $k$ are in the same social position within a dimension, the distance between them is $x_{jk} = 1$. Colleagues at work, classmates at school, and members of the same country club all have a social distance of 1. Measuring the distance between the individuals $i$ and $j$ in figure 1 requires going up two branches in the tree to reach their common ancestor. Starting from a social distance of 1, these two steps increase their social distance to $x_{ij} = 3$.

The same ultrametric scale that is used to measure social distance between individuals within a dimension can also be used to measure the distance between a single individual's social positions across different dimensions. Figure 2 shows that in $d_2$, individual $i$ has a social distance of $y_i^{12} = 3$ from her position in $d_1$. One noteworthy feature of the ultrametric measure of social distance is that every position on the opposite "side" of the social tree is equally far away (since the maximum distance is the uppermost fork). Thus, individuals in position 5 are at a distance of 4 from individuals in positions 1–4. Or, in other words, while members of the labor positions have a fine-grained sense of their social distances from their fellow labor workers, all the members of management are equally "far away."[6]

[5] The schema shown in figs. 1 and 2 can also be represented with linear distances instead of ultrametric distances. However, Watts et al. (2002) show that the ultrametric formalization of social distance provides a good fit with data on social distance in social networks and matches the predictions of Kleinberg (2000) on the effects of homophily on network navigability.

[6] Following Blau and Schwartz (1984), these social distances represent perceived social distances, which "refer only to those differences among people in terms of which they themselves make social distinctions in their relations with one another" (p. 192). Similarly, Watts et al. (2002, p. 1303) argue that this approach to modeling social distance preserves the intuitive way in which actors make "cognitive divisions" that "break down, or partition, the world hierarchically into a series of layers . . . of increasingly specific groups." My formal representation of social distance is based on current cog-

## Homophily

"Homophily," in Blau and Schwartz's terms, refers to the likelihood that people in similar social positions will tend to be socially connected. Homophily can naturally emerge as a result of "choice homophily": individuals' preferences to form ties with similarly positioned alters (Blau and Schwartz 1984; McPherson et al. 2001; Centola and van de Rijt 2015). At the same time, people may also choose to form social ties that are more diverse. For instance, in the work setting, while people interact disproportionately with others who have jobs of the same rank, those people might not be the most interesting to us. Some executives may play pickup basketball during lunch with workers from the mailroom, or bike messengers may meet middle managers outside during smoke breaks and find that they are enjoyable company. However, even if individuals prefer to have diverse social contacts, low permeability of institutional barriers (McPherson and Smith-Lovin 1987), or social disapprobation for having friends from outside the neighborhood or office community, can force homophilous tie formation even in the absence of homophilous preferences (McPherson and Smith-Lovin 1987; McPherson et al. 2001). Consequently, in each dimension, the social ties that form are not just a function of the individuals who occupy each social position but a result of both individual and structural forces that constrain tie formation.

I represent the level of homophily in the process of tie formation as a parameter that controls each individual's "radius of search" in his or her selection of social contacts.[7] If homophily is strong, the radius is very restricted: individuals search only within a social distance of 1 (i.e., within their own social position) to make new ties. If homophily is weaker, the radius increases: for example, radius 2 allows individuals to randomly make friends across all social positions within distance $x_{ij} \leq 2$. If there is no homophily, ties are made at random across all positions.

nitive psychological models, which posit a negative-exponential relationship between social-cultural distance and perceived similarity across diverse domains of social judgment (Shepard 1987; Tenenbaum and Griffiths 2002). Building on previous work, my model is intended to be general across a variety of social domains and settings in which actors discriminate social distances on the basis of increasing levels of similarity and difference (such as employment, residential neighborhood, and aesthetic sophistication). However, to test the robustness of my model, I also explored alternative implementations of social distance (e.g., using linear distances) and found that, consistent with the results presented here, in the alternative model, both network structure and diffusion depend on high levels of homophily and consolidation.

[7] This radius can be interpreted "agentially," in terms of individuals' predilections for homophilous or heterophilous tie formation, or "structurally," in terms of variations in the strength of institutional and social constraints on contact rates between the members of diverse social positions.

### Heterogeneity, Complexity, and Group Size

While consolidation and homophily are the primary theoretical parameters under consideration here, there are three other parameters that are important for evaluating the robustness of our results. The first of these is "heterogeneity." Put simply, this is the number of social positions within each dimension of social life.[8] When there is very low heterogeneity, opportunities for social differentiation are minimal. At the extreme, if there is only one group in each dimension of social life (e.g., in Shaker societies, where every dimension was populated with universal participation in a single position), then there are no opportunities for social differentiation, homophily, or consolidation. It is, quite simply, a homogeneous population. As heterogeneity increases, opportunities for differentiation increase (e.g., more possible jobs to work at, more schools to attend, more clubs to join, etc.), making homophily noticeably distinct from random interactions and making consolidation very different from intersecting social circles (Blau 1977; Blau and Schwartz 1984).

"Complexity" is the complement of heterogeneity. While heterogeneity is the number of different social positions per dimension, complexity is the total number of dimensions. As complexity increases, the role of consolidation becomes more pronounced. If there are only two dimensions (e.g., home and work), the potential effects of consolidation are limited since the number of possible locations for tie formation is so small. However, as there are more dimensions to social life (e.g., neighborhood, work, lodge memberships, sports clubs, book clubs, church groups, etc.), the effects of consolidation can produce much greater segregation in society, significantly changing the opportunities for crosscutting ties to be created through the diversity of social dimensions. (More detailed descriptions of both heterogeneity and complexity are provided in Blau and Schwartz [1984].)

Finally, "group size" is the number of people located in each social position. The size of the total population being studied ($N$) is the product of group size and heterogeneity: The number of people in each position times the total number of positions gives the full size of the population. Many previous studies have examined the effects of heterogeneity in group size on homophily in social networks (Blau 1977; Skvoretz 1983; Blau and Schwartz 1984; Sampson 1984; McPherson and Smith-Lovin 1987; McPherson et al. 1992; Reagans, Zuckerman, and McEvily 2004).[9] However, for the purposes

[8] See the appendix for details on the use of these parameters.

[9] The expected frequency of tie formation between groups under random mixing (i.e., no choice homophily) depends on the relative sizes of those groups. For example, if there are 10 boys and 90 girls and everyone has 10 social ties, then "random mixing" predicts that each girl will have nine "homophilous ties" and one "heterophilous tie," while the boys will have the opposite. But, despite girls having many more "same-type" ties than boys, they do not exhibit homophilous bias in their social ties if their networks do not

of the present study, group size is treated as a constant (uniform across all social positions), with the same number of individuals in every social position across all dimensions (as shown in figs. 1 and 2). This model obviously greatly simplifies the complexity of the real world. However, in doing so, it provides an analytical tool for identifying the effects of consolidation and homophily on the emergence of network topologies and the resulting implications for the dynamics of social diffusion.

Together, the full suite of parameters for the model is as follows:

- $\alpha$ = homophily, ranging from $-1$ to 3.
- $\beta$ = consolidation, ranging from $-1$ to 3.
- $H$ = heterogeneity (number of social positions), ranging from 1 to infinity.
- $D$ = complexity (number of dimensions), ranging from 1 to infinity.
- $G$ = group size (number of people in each position), ranging from 1 to infinity.
- $N$ = population size = $H \times G$.
- $Z$ = average degree (number of ties for each person), ranging from 1 to $N - 1$.

## The Process of Network Formation

The model is initialized with $N$ individuals, each assigned to one social position in each of the $D$ dimensions.[10] Each individual is assigned to a random position within the first dimension ($d_1$). Individuals' positions in subsequent dimensions are determined by the consolidation parameter, which is in the range $-1 \leq \beta \leq 3$. When consolidation is set to $\beta \approx -1$ (i.e., zero consolidation), an individual's social position is chosen at random in each dimension, irrespective of her positions in other dimensions. This means that in a population with eight social positions in every dimension (i.e., heterogeneity = 8), the typical distance between an individual $i$'s positions across two dimensions $a$ and $b$ is 4 (i.e., $y_i^{ab} = 4$), which is the maximum social distance between positions.

deviate from the random mixing assumption. Thus, the meaningfulness of intergroup ties in a population depends on the relative group sizes and whether tie frequencies deviate from the predictions of random mixing. Detailed and interesting work on this problem has been done by Blau (1977), Skvoretz (1983), Blau and Schwartz (1984), Sampson (1984), Blum (1985), McPherson and Smith-Lovin (1987), Knottnerus and Guan (1997), Reagans (2005), and others. Because of the complexity of the present model, the number of individuals is held constant across social positions in order to make the analysis tractable. Additional analyses indicate that my results are robust to varying assumptions about the distribution of group sizes; however, expanding the model to explore the effects of heterogeneity in groups size on the dynamics of diffusion provides an interesting direction for future research.

[10] The appendix provides complete technical details for the model.

As consolidation increases, each person's position becomes more correlated across dimensions. A value of $\beta \approx 0$ (low consolidation) means that given a person's position in one dimension of social life, we can predict her location in other dimensions with 25% accuracy. If heterogeneity = 8, the typical distance between an individual's positions across two dimensions $a$ and $b$ is approximately 3, that is, $y_i^{ab} = 3$. Moderate consolidation (a value of $\beta \approx 1$) means that we can predict a person's location across all dimensions with 66% accuracy just from knowing a single social position. The typical distance between an individual's positions is thus reduced to approximately 1.5, that is, $y_i^{ab} = 1.5$. Very high consolidation (a value of $\beta \approx 2.5$) means that we can predict with 90% accuracy what position a person will be located in across all dimensions just from knowing her social position in one dimension. This reduces the expected distance between positions to approximately 1; that is, people are typically in the same position in every dimension.

Once every member of the population has been assigned social positions in all $D$ dimensions, network formation begins by (i) selecting an individual $i$ at random, (ii) selecting a dimension $d$ at random, and (iii) selecting a social position that falls within distance $x_{ij}$ from $i$'s own social position. The distance $x_{ij}$ reflects the level of homophily in tie formation. Just as with consolidation, the parameter for homophily, $\alpha$, ranges over the interval $-1 \leq \alpha \leq 3$, ranging from random choice (no homophily) to near-perfect homophily in tie formation. When homophily is $\alpha \approx -1$, ties are made at random, irrespective of actor $i$'s social position (i.e., $x_{ij} = 4$). A low level of homophily ($\alpha \approx 0$) corresponds to modest restrictions on which social positions people draw their social ties from (i.e., $x_{ij} = 3$). Moderate levels of homophily ($\alpha \approx 1$) restrict tie formation to adjacent social positions ($x_{ij} = 1.5$). And very high homophily ($\alpha \approx 2.5$) constrains almost all social contact to in-group tie formation ($x_{ij} = 1$).

Once $i$ selects a social position that falls within the distance $x_{ij}$, she then selects an individual $j$ at random from that position, and they create a social tie. The homophily parameter is the same for all actors in the model. The number of ties that each actor has is controlled by the "degree" parameter, $Z$. The tie formation process iterates until, on average, each individual has $Z$ ties. Ties are treated as undirected, which means that every relationship is symmetric. In order to identify the emergent properties of network topology and how they affect diffusion processes, I study sparse networks such that $Z \ll N$ (Watts and Strogatz 1998; Watts 1999).

## TOPOLOGICAL FEATURES OF EMERGENT NETWORKS

To begin our exploration of network formation, I assume fixed values of heterogeneity, complexity, group size, and average degree. I then conduct computational "experiments" in which I systematically alter the levels of consolidation and homophily. Subsequently, I examine the robustness of the results across variations in each of the other parameters, including heterogeneity ($H$), complexity ($D$), and group size ($G$).[11] With the network formation algorithm described above, structurally embedded individuals make social ties, which collectively produce emergent social networks that can be analyzed according to their topological features. Starting from the theoretical limiting cases of random assignment to social positions (zero consolidation) and random choice of social ties (zero homophily), I independently increase the levels of consolidation and homophily in the population.[12] Table 1 reports the topological and social features of the resulting social networks,[13] and figure 3 shows corresponding graphical representations of the network topologies.

Each column in table 1 represents a different level of social consolidation, ranging from purely random assignment of social positions (zero consolidation) in column 1 to strong correlations between an individual's positions across dimensions (very high consolidation) in column 4. Respectively, the four lines in each row show the corresponding values for homophily, ranging from randomly chosen ties (zero homophily) in line 1 to strong preferences for in-group tie formation (very high levels of homophily) in line 4. Figure 3 shows the network structures that correspond to each of the 16 possible combinations represented by each row of table 1.[14]

The first row in table 1 reports the clustering coefficient ($CC$), which is the fraction of each individual's neighbors who are neighbors with each other (i.e., closed "triads" in the social network; Watts and Strogatz 1998; Newman 2000). The clustering coefficient is normalized by calculating the limiting case clustering coefficient ($CC_0$) for a "randomized" version of the same network. This is the theoretical lower bound for clustering in a social network (Watts 1999). The normalized measure ($CC/CC_0$) indicates the relative increase in the level of social clustering above the lower theoretical limit (Watts and Strogatz 1998). This measure provides a general estimate of the cohesiveness of individuals' social ties within the network (Granovetter 1973; Moody and White 2003; Watts 2003). Empirical social

[11] Consistent with Blau and Schwartz (1984), I find that changes to the distribution of resources (i.e., the level of status inequality) do not affect the structural properties of the emergent social networks.

[12] I also explored the effects of heterophily on network formation by extending the parameter space into the negative domain. I found that the results from this extension of the parameter space did not differ qualitatively from those of random tie formation.

[13] The results in table 1 report ensemble averages from 100 independent runs of the model.

[14] The four cells that are marked with an asterisk in each row of table 1 correspond to the four panels shown with a gray background in fig. 3.

TABLE 1
Topological Properties of Emergent Social Networks
($H = 32$, $D = 10$, $Z = 10$, $G = 100$)

| Homophily | Consolidation: Random | Low | Moderate | High |
|---|---|---|---|---|
| Clustering coefficient/$CC_0$: | | | | |
| Random | 1.00 | 1.01 | 1.0469 | 1.1257 |
| Low | 1.2536 | 3.4203 | 7.9203* | 9.705* |
| Moderate | 1.4060 | 5.8735 | 22.9598* | 33.2815 |
| High | 1.4786 | 5.9736 | 28.0952* | 42.016 |
| Path length/$L_0$: | | | | |
| Random | 1.00 | 1.00 | 1.00 | 1.00 |
| Low | 1.00 | 1.03 | 1.10* | 1.13* |
| Moderate | 1.00 | 1.06 | 1.36* | 2.16 |
| High | 1.00 | 1.06 | 1.39* | Infinite |
| Bridge width: | | | | |
| Random | .6807 | .6967 | .6854 | .6938 |
| Low | .6948 | .8978 | .9616* | .9597* |
| Moderate | .7138 | .9406 | .9897* | .9925 |
| High | .6992 | .9410 | .9919* | .9963 |

Note.—Network clustering, path length, and bridge width are shown for each combination of homophily and consolidation. Columns show consolidation increasing from random to high levels, and the lines in each row show the corresponding increase in homophily, from random to high levels.
* Network values where social diffusion is successful.

networks typically have many times more clustering than what is found in corresponding random networks (Watts 1999; Newman 2000; Newman et al. 2006), indicating that people tend to create and sustain clustered ties in their relationships.

Line 1 in the first row in table 1 shows that without homophily (i.e., when ties are formed at random), network clustering approximates a random network for all values of consolidation. Even when consolidation is very high (col. 4), overall clustering approximates the theoretical lower limit of a random network. Similarly, the first cell of row 1 shows that when consolidation is held constant at zero, increasing homophily (lines 1–4) also has a negligible effect on the level of clustering in the network. As suggested by Skvoretz (1983) and Blau and Schwartz (1984), homophily and consolidation are highly interdependent. Increasing one without any contribution from the other results in no meaningful change in the level of clustering.

However, the second line in row 1 shows that once homophily increases even modestly, increasing consolidation has a marked impact on the emergence of clustering in the network. Moving from cell 1 (zero consolidation) to cell 4 (very high consolidation) increases the level of clustering in the network nearly eightfold. Moreover, lines 3 and 4 in row 1 show that this

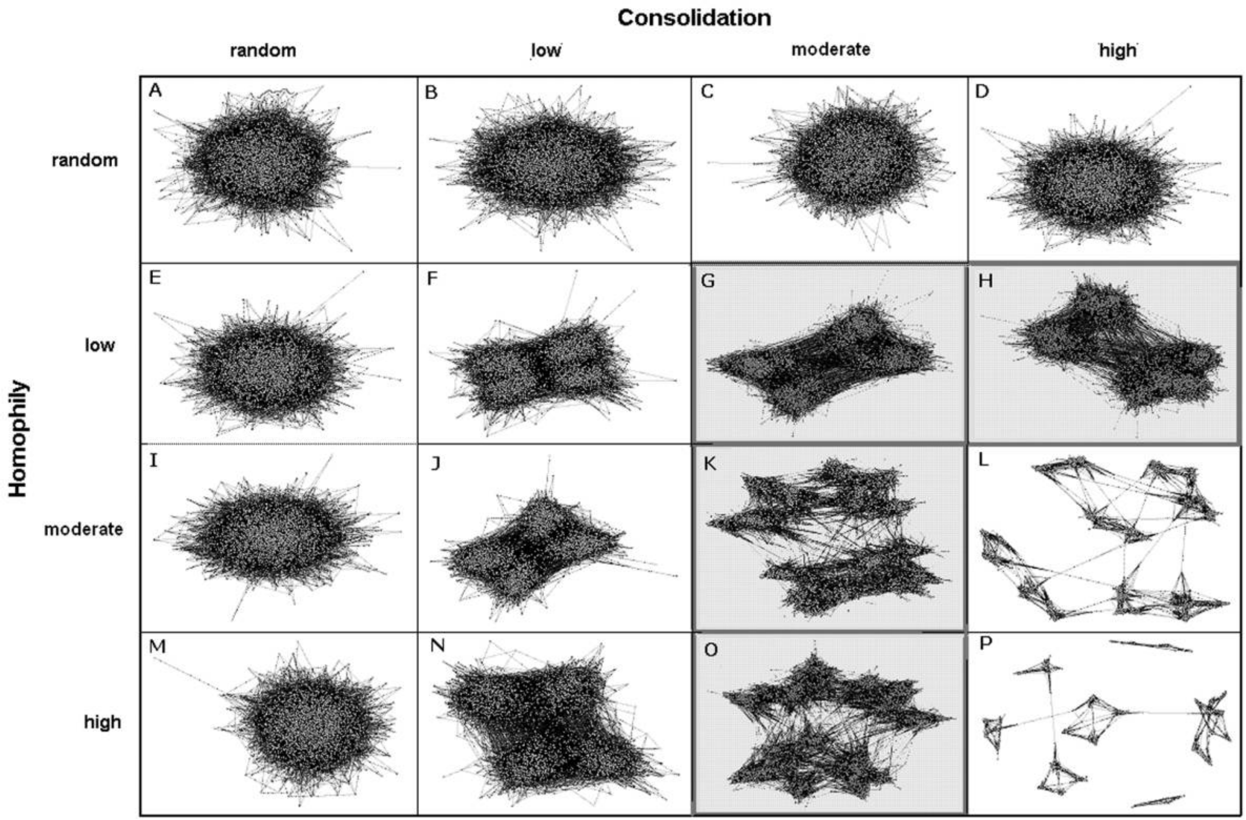


FIG. 3.—Emergent network structures ($H = 32$, $D = 10$, $Z = 10$, $G = 100$). Social network topologies emerge from the interaction of homophily and consolidation. Columns show consolidation increasing from random to high levels, and rows show homophily increasing from random to high levels. Panels shown in gray correspond to topologies for which social diffusion is successful.

effect increases with higher levels of homophily. When homophily is very strong (line 4), going from zero consolidation to high consolidation causes network clustering to jump by a factor of 28. In other words, the greater the tendency for individuals to interact with similar others, the more pronounced the effect of social structures that consolidate people's opportunities for interaction on the level of network clustering in the emergent social topology.

Each cell in row 1 shows a similar result for the effects of increasing levels of homophily, assuming a fixed level of consolidation; that is, each of the cells in row 1 reports the effects of increasing homophily while maintaining a fixed level of social consolidation. When one goes from line 1 to line 4, the effects of homophily become more pronounced as social consolidation increases (going from cell 1 to cell 4). In cell 4, the increase from zero homophily to very high levels of homophily results in a 37-fold increase in the cohesiveness of the network structure. Thus, consistent with previous work (Blau 1977; McPherson and Smith-Lovin 1987), the effects of homophilous preferences on the structure of the emergent social network depend on the institutional constraints that consolidate a society.

Row 2 of table 1 shows the normalized characteristic path length ($L/L_0$) of the emergent social networks. Characteristic path length, $L$, reports the average number of social links that need to be traversed in order to travel between any two nodes in the network (Watts 1999). Row 2 shows the normalized value of $L/L_0$ for increasing values of both homophily and consolidation. A topology with a random structure and no clustering is a "small world," which means that any node can be reached from any other node in just a few steps (Newman 2000).[15] Networks with zero consolidation and zero homophily are random networks with minimal clustering and minimal path length. As we saw for clustering, there is a similar interdependence between homophily and consolidation on the characteristic path length of the network. Increasing one parameter (e.g., homophily) while leaving the other (e.g., consolidation) at zero produces no significant changes in the network's path length; it still approximates a random network. However, increasing either homophily or consolidation to moderate levels allows increases in the other parameter to significantly increase the path length of the network by introducing clustering into the ties, thereby increasing the number of steps that need to be traversed in order for a contagion to travel across the population. As shown in Watts and Strogatz (1998) and Watts (1999), characteristic path length is a key network parameter for understanding the dynamics of social diffusion: as path length decreases, so do diffusion times.

[15] These are distinct from small-world networks, which have high levels of clustering and a low characteristic path length (Watts and Strogatz 1998).

Consequently, the effect of increasing or decreasing consolidation on the topological structure of a society's social network has direct implications for the dynamics of social diffusion. The more intersecting that a society is—the less clustering in the network and the smaller the characteristic path length—the easier it should be for collective behaviors to spread (Granovetter 1973; Watts and Strogatz 1998). Coordination on norms, the spread of social movements, and the diffusion of cultural practices should all be easier as consolidation is reduced (Granovetter 1973; Blau and Schwartz 1984; Macy 1990). The institutional and demographic features of a society that control social consolidation can thus unwittingly determine a society's susceptibility to emergent fads, widespread cultural coordination, and even the success of social cooperation (Granovetter 1973, 1978; Watts 2002; Centola and Macy 2007).

## HOW HOMOPHILY AND CONSOLIDATION AFFECT SOCIAL DIFFUSION

As Blau and Schwartz (1984) observe, highly consolidated populations can become balkanized into micro regimes of politically, economically, and racially homogeneous groups. This is the nadir of social integration, which implies that there will be low levels of social diffusion since behaviors and practices cannot spread between segregated groups (e.g., panels *L* and *P* in fig. 3). By contrast, when consolidation is minimal, there are abundant crosscutting ties between highly diverse individuals, with little or no structural boundaries to inhibit social interactions (e.g., panels *A*, *E*, *I*, and *M*). The major implication of Blau and Schwartz's theory of social structure is that reducing social consolidation in a population creates more crosscutting ties in the social network, thereby promoting social diffusion across the society.

The main finding of this study is that while, on the one hand, excessive consolidation can fracture a social network (as in panels *L* and *P*), on the other hand, eliminating consolidation entirely can in fact erode the social bridges that maintain the integrity of the social network. Reducing consolidation too far can eliminate the group structures that provide social support for collective behaviors, which allow behaviors to be transmitted across diverse parts of the population. Surprisingly, I find that the social groups that emerge when structural consolidation actively constrains the process of network formation are in fact necessary in order to produce network topologies that can support the dynamics of social integration.

I measure social integration in terms of the successful spread of a cultural practice, behavioral norm, or collective belief across the population. The diffusion process is initiated by activating a randomly chosen "seed neighborhood" of early adopters (consisting of an individual and her $z$ immediate "neighbors" in the social network). Individuals are then chosen asynchronously, at random, to evaluate whether to adopt the behavior.[16] Unlike the diffusion of a "simple contagion," such as a disease or a piece of information, which requires only a single contact for transmission, social norms and cultural practices are typically "complex contagions," which require social reinforcement from multiple sources in order to be transmitted (Granovetter 1978; Boyd and Richerson 1985; Castro and Toro 2004; Centola and Macy 2007). For example, the presence or absence of normative validation can affect an individual's decision to join a social movement (Marwell and Oliver 1993), wear a new fashion (Grindereng 1967), or go on strike (Klandermans 1988). The reason is not only that more contact with adopters provides reinforcing signals that a behavior is validated but also that having more contact with adopters also reduces contact with peers who would pressure someone not to adopt. Indeed, as a behavior becomes more normatively validated, nonadoption can also become less validated.

Additionally, strategic complementarity (Granovetter 1978; Schelling 1978; Marwell and Oliver 1993)—the increasing rewards created by a "critical mass" of friends and contacts who adopt a behavior—and emotional contagion—the excitement generated by sharing a common practice or activity with others—are also powerful reasons why social reinforcement plays such an important role in the spread of new behaviors (Collins 1993). For the widespread integration of cultural practices (Axelrod 1997; Centola et al. 2007), cooperative behaviors (Axelrod 1984; Kim and Bearman 1997; Centola and Macy 2007), and social norms (Centola, Willer, and Macy 2005), normative validation, complementarity, and emotional contagion (Centola and Macy 2007) may all be salient reasons why successful diffusion depends on reinforcing social signals.

In what follows, I study the dynamics of social integration as dependent on social reinforcement, such that each individual requires contact with at least two sources before he or she will be willing to adopt a new behavior (Centola, Eguiluz, and Macy 2007). More formally, every individual in the population is assigned a threshold for adoption ($T = r/z$), whereby he or she will adopt the new practice, behavior, or belief only if there is sufficient reinforcement, $r$, from his or her neighbors. I study the dynamics of social integration by setting $r = 2$.[17] Thus, the behavior spreads to new people

[16] Asynchronous updating in random order and without replacement eliminates potential order effects and guarantees that every node is updated in every round of decision making, defined as $N$ time-steps.

[17] Centola and Macy (2007) show that the qualitative results that hold for $r = 2$ are robust for a wide range of assumptions about deterministic and probabilistic threshold distributions.

only if at least two of their neighbors have adopted it; this is the minimum level of social reinforcement for a complex contagion (Centola and Macy 2007).[18]

The success of social diffusion is measured in terms of the fraction of the population that adopts the collective behavior. Figure 4 shows the success of social diffusion (along the $z$-axis) for increasing values of social consolidation (along the $x$-axis) and homophily (along the $y$-axis). Along the $x$-axis, social consolidation increases from zero correlation ($\beta = -1$) to nearly perfect correlation ($\beta = 3$). Similarly, along the $y$-axis, homophily in individual tie formation increases from zero homophily ($\alpha = -1$) to nearly perfect homophily ($\alpha = 3$). The values along the $z$-axis indicate the fraction of the population that is reached by the diffusion process (averaged over 100 runs of the model). The results in figure 4 are for group sizes of $G = 100$, where the number of groups is $H = B^{L-1} = 2^5 = 32$, the number of dimensions is $D = 10$, and the average number of ties per person is $Z = 10$. The total population size ($N$) is given by the number of groups times the size of each group: $N = B^{L-1} \times G = 3{,}200$.

Figure 4 shows that diffusion succeeds for only a specific range of values of homophily and consolidation. (The network topologies corresponding to these values of homophily and consolidation are indicated by the starred elements of table 1 and those in fig. 3 shown with gray backgrounds.) When ties are not homophilous, diffusion fails entirely regardless of the level of social consolidation. This can be seen by tracing the gridline corresponding to $\alpha = -1$ along the increasing values of consolidation. Throughout the entire range of values along the $x$-axis, social diffusion never increases. Correspondingly, when social positions across dimensions are randomly assigned ($\beta = -1$), increasing the level of homophily (i.e., following the $\beta = -1$ gridline along the $y$-axis) does not create sufficient changes to the social structure to improve the dynamics of diffusion.[19] Even when homophily is very strong ($\alpha > 2$), the lack of consolidation makes social tie formation an essentially random process. Consequently, if social structures provide too few constraints on who has contact with whom, the resulting social network will "behave" identically to one in which everyone interacts at random. The effects of low levels of homophily cannot be distinguished from the effects of low levels of structural consolidation (cf. Mc-

[18] An additional factor affecting social reinforcement is the similarity between contacts, where social contacts who are more similar to one another may be more likely to influence behavior change (Axelrod 1997; Centola 2011). I exclude this assumption from the current model in order to provide a conservative test of whether increases in homophily and consolidation can facilitate diffusion. Alternative realizations of the model that include this assumption show that the results exhibit the same qualitative dynamics as those presented here.

[19] However, there are changes to the network (cf. Watts et al. 2002).

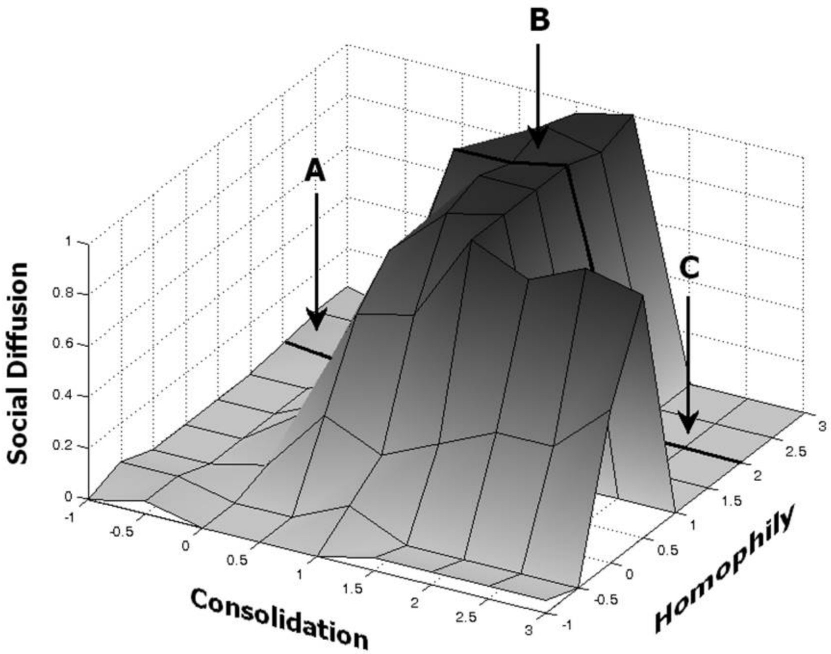


FIG. 4.—Successful diffusion by homophily and consolidation ($H = 32, D = 10, Z = 10, G = 100$, averaged over 100 realizations). Fraction of the population that adopted the collective behavior is shown along the $z$-axis. The $x$-axis shows consolidation increasing from random to perfect consolidation. The $y$-axis shows homophily increasing from random to perfect homophily. For both low and very high values of consolidation, diffusion fails. For moderate to high levels of consolidation, social diffusion is successful. The points $A$, $B$, and $C$ correspond to the success of diffusion with random ($\beta = -0.7$), moderate ($\beta = 1.3$), and very high ($\beta = 2.5$) levels of consolidation and high ($\alpha = 2$) levels of homophily.

Pherson et al. 2001): Either of these social forces can "trump" the other one, eliminating the social infrastructure necessary to support the dynamics of social diffusion.

The key to these dynamics is shown in row 3 of table 1, which reports the fraction of neighborhoods in the network with at least one "wide bridge" to another neighborhood, where "bridge width" is measured as the number of ties between an individual's neighbors and the focal node of a nonadjacent neighborhood (Centola and Macy 2007). The overall fraction of wide bridges reports the number of neighborhoods that can be directly "activated" by a nearby neighborhood. In row 3, line 1 and cell 1 indicate, respectively, that when either homophily or consolidation is zero, only 68%–70% of the population has a wide bridge sufficient to transmit a new behavior, where adoption requires minimal social reinforcement ($r = 2$). As homophily and consolidation increase, so does the formation of wide bridges, which creates more pathways for social reinforcement (Centola, Eguiluz, and Macy 2007).

The dynamics of social diffusion depend on wide bridges formed by overlapping patterns of social relations. These emerge from the interaction of the social and institutional forces of homophily and consolidation. While this is consistent with Blau and Schwartz's (1984) finding that consolidation can dramatically affect the structure of social networks, it also shows something unanticipated: Moderate to high levels of consolidation can be necessary for creating social structures that support the dynamics of widespread social integration. Without social consolidation, the collective body of social ties that emerges from even highly homophilous social interactions will be incapable of supporting the diffusion of complex social behaviors.

To illustrate this, figure 5 shows the success of social diffusion for a single, high value of homophily, $\alpha = 2$, as a function of increasing values of social consolidation ($-1 \leq \beta \leq 3$). Figure 5 corresponds to the darkened gridline in figure 4, where the homophily parameter is held constant at $\alpha = 2$. This is the value of $\alpha$ that Watts et al. (2002) find to be a good approximation of empirical levels of homophily in social network formation.[20] In both figures, the locations $A$, $B$, and $C$ correspond, respectively, to diffusion with zero consolidation, moderate consolidation, and very high consolidation. Even with a high level of homophily, moderate levels of consolidation are required in order for social diffusion to succeed. Following the progression from low to high consolidation, figure 6 shows the underlying topologies that correspond to locations $A$, $B$, and $C$ in both figures 4 and 5.

Panel $A$ of figure 6 shows the social network of a population with high levels of homophily ($\alpha = 2$) but with zero consolidation ($\beta = -1$), corresponding to position $A$ in figures 4 and 5. The level of clustering in this network is very low ($CC/CC_0 = 1.47$), as is the fraction of the population that can be reached by wide bridges (fraction wide bridges = 0.69). For the network shown in panel $A$, 30% of neighborhoods cannot propagate complex contagions to nearby neighbors. As indicated by the value of the $z$-axis in figure 4 (labeled $A$), diffusion is unsuccessful in this network (less than 1% adopt the behavior) because of the lack of social structure to support reinforcement for behavioral adoption (Centola and Macy 2007).

[20] This value corresponds with Kleinberg's $\alpha$ (Kleinberg 2000) and provides a good estimate of the conditions under which social networks are searchable. This value is calculated by Watts et al. (2002) using a model similar to the one presented here.

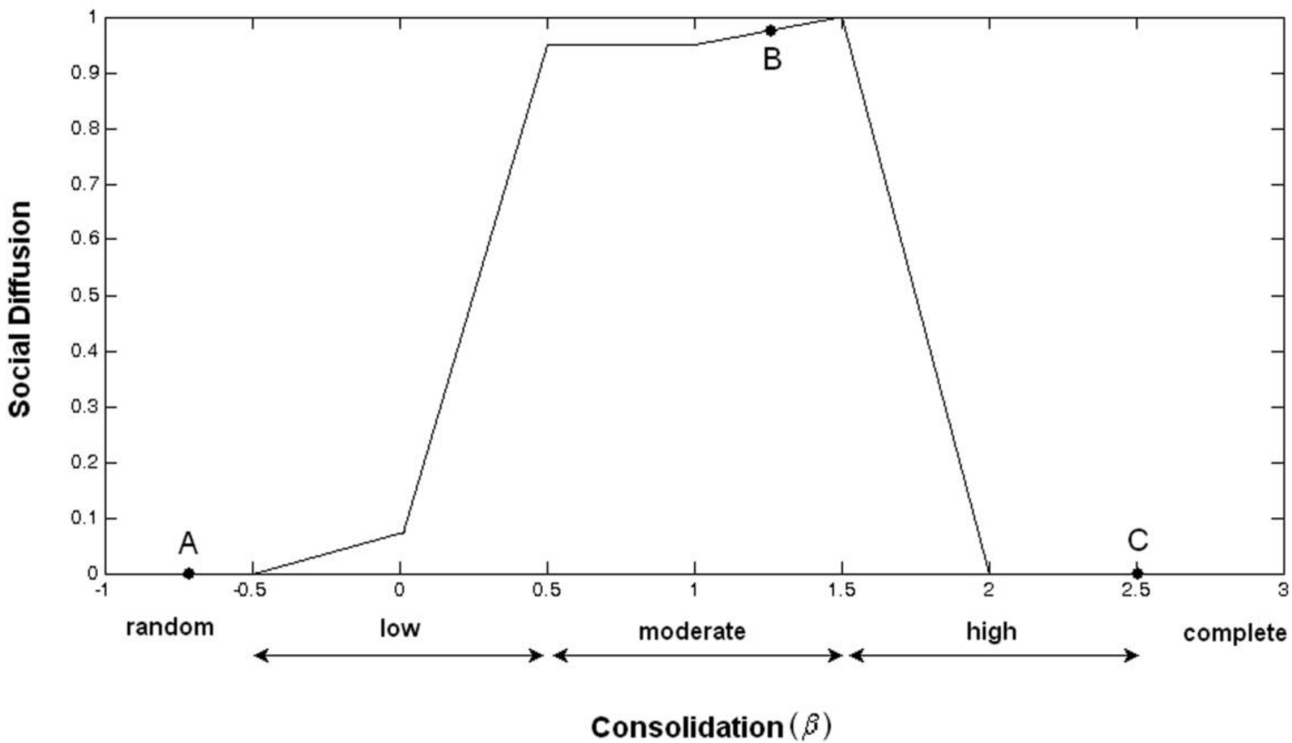


Fig. 5.—Successful diffusion by consolidation ($H = 32$, $D = 10$, $Z = 10$, $G = 100$, $\alpha = 2$, averaged over 100 realizations). Fraction of the population that adopts the collective behavior is shown along the $y$-axis, for increasing levels of consolidation, from random to perfect, along the $x$-axis. For both low and very high values of consolidation, diffusion fails. For moderate to high levels of consolidation, social diffusion is successful. The points $A$, $B$, and $C$ correspond to the success of diffusion with random ($\beta = -0.7$), moderate ($\beta = 1.3$), and very high ($\beta = 2.5$) levels of consolidation.

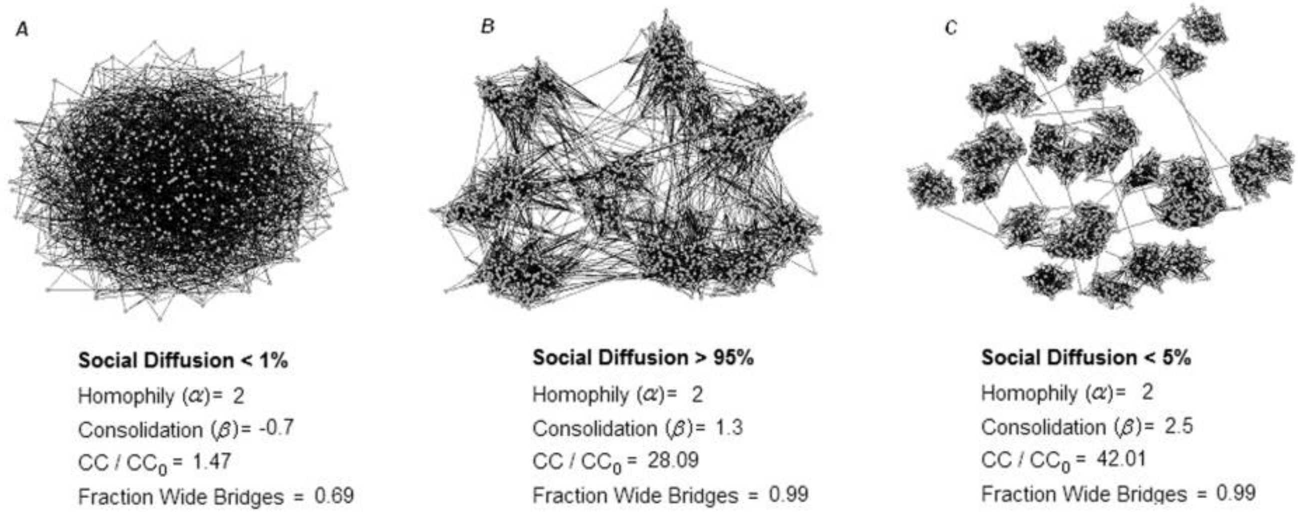


FIG. 6.—Changes in network structure with increasing consolidation ($H = 32, D = 10, Z = 10, G = 100, \alpha = 2$). Network topologies correspond to points $A$, $B$, and $C$ in figures 4 and 5. For zero consolidation (panel $A$), the network is unstructured and cannot support social diffusion. Moderate levels of consolidation (panel $B$) produce overlapping group structures and wide bridges, which support the spread of the behavior across the population. For very high levels of consolidation (panel $C$), the network is fragmented and diffusion fails.

Panel *B* in figure 6 shows the social network structure that emerges from the same amount of homophily ($\alpha = 2$), but with much higher levels of social consolidation ($\beta = 1.3$). This social network has 17 times greater clustering ($CC/CC_0 = 28.09$) than the network in panel *A*, and 99% of neighborhoods can be reached by bridges wide enough to transmit a minimally complex contagion. Consequently, as indicated by the corresponding arrow in figure 4 (labeled *B*), diffusion is much more successful in this social structure, resulting in adoption by over 95% of the population.

The statistical differences between the networks shown in panels *A* and *B* of figure 6 also present a striking visual difference between them. While the network in panel *A* is relatively undifferentiated from region to region, the network in panel *B* has clearly defined local clusters. In panel *B*, distinct regions of clustered social activity are held together by densely overlapping ties between groups. In the middle of the figure, we can discern two groups that play a "brokering" role between the other, more socially distant groups. As seen in panels *K* and *O* in figure 3, these brokering groups typically emerge when the levels of homophily and consolidation are within the range shown in panel *B* of figure 6, where diffusion is highly successful. This suggests that the emergence of wide bridges across a large population may depend on the formation of group structures that maintain overlapping relationships between remote regions of the social space.

As observed by Granovetter (1973) and others (Burt 1992; Watts 1999), when individuals act as information brokers, they can unite distant regions of social space. But they cannot provide overlapping ties and wide bridges that bind those regions together with shared customs and practices (Centola and Macy 2007). The figure in panel *B* suggests that social diffusion in large, complex societies may depend on emergent group-level brokering structures that bind socially remote groups together through their overlapping ties with more socially "intermediate" groups.

Figure 6 also shows that the formation of group structures can go too far. At the extreme, densely knit group structures can also become so clustered as to prevent meaningful contact with nongroup members. Panel *C* in figure 6 shows the social network that emerges with very high levels of social consolidation ($\beta = 2.5$), such that people's social positions are nearly perfectly correlated across all dimensions. In panel *C*, clustering is 28 times greater than it is in panel *A* ($CC/CC_0 = 42.016$), and almost every near neighbor can be reached by a wide bridge (fraction wide bridges = .9963). The downside of these clustered ties and wide bridges is that overlapping memberships across groups have been lost, leaving only weak ties across socially isolated islands of normative autonomy (Centola et al. 2007). Within any given group, there is easy convergence on a custom or belief, but across groups, there are no wide bridges to support social influence, eliminating the possibility of widespread social integration. As shown by the corresponding arrow in figure 4 (labeled $C$), because diffusion is restricted to small clusters, the global spread of a new custom or norm is limited to less than 5% of the total population.[21]

## ROBUSTNESS ACROSS COMPLEXITY, HETEROGENEITY, GROUP SIZE, AND STATUS

My results demonstrate the importance of consolidation and homophily for structuring social relations to support the integration of shared customs and beliefs across a population. Going a step further, I investigated the robustness of these results by increasing and decreasing the number of dimensions (i.e., complexity, $2 \leq D \leq 20$), changing the number of groups (i.e., heterogeneity, $8 \leq B^{L-1} \leq 64$), and changing the sizes of the groups (i.e., $50 \leq G \leq 400$). Throughout these experiments, I kept the size of the population constant ($N = 3{,}200$) and the density of ties constant ($Z = 10$) so as not to conflate network density effects with effects of the structural parameters.[22] There were only two qualitative deviations from the results presented above. First, when there is a very low number of dimensions of social life ($D \leq 3$), social integration succeeds even with zero consolidation ($\beta = -1$). The reason is that when there are only a few dimensions, there is not enough complexity in the social structure to prevent ties from forming overlapping clusters. Work ties, friendship ties, and neighborhood ties all wind up overlapping with the same subsets of people, which allows wide bridges to form despite the lack of consolidation. As the number of dimensions increases slightly above the bare minimum ($D > 3$), the diversity of domains of social interaction makes it possible for actors to form socially

[21] A pattern similar to the one shown in panels $A$, $B$, and $C$ of fig. 6 can also be found by fixing the consolidation parameter at $\beta = 2$ and increasing homophily from $\alpha = -1$ to $\alpha = 3$. As shown in fig. 4, when consolidation is high ($\beta = 2$) but social ties are random ($\alpha = -1$), networks are too diffuse to support the spread of social behavior (diffusion < 1%). As ties become increasingly homophilous ($\alpha \geq 0.5$), emergent clusters broker relationships between social groups, facilitating the dynamics of social integration (diffusion > 95%). But as homophily increases ($\alpha > 1$), the network forms into weakly connected clusters of locally homogeneous, socially segregated groups, once again preventing integration (diffusion < 5%).

[22] In order to identify the effects of structural parameters on social diffusion, my study is restricted to large, sparse networks ($Z \ll N$) such that significant changes in network topology can occur (Watts and Strogatz 1998; Watts 1999; Newman 2000; Watts et al. 2002; Centola and Macy 2007). Related empirical and theoretical work has focused on much more sparse networks (0.0003 < density < 0.0009). The results presented here demonstrate the robustness of my results for networks ranging up to density ~0.003, and additional robustness tests have generalized the results to networks with density up to ~0.006. Consistent with earlier work on network structure (Watts 1999), for higher densities I find no effects of structural variation on collective behavior.

distant ties. Once this is possible, social consolidation is necessary for the formation of local clusters and wide bridges.

The other deviation occurred with increasing local group size, $G$. But instead of diffusion becoming less dependent on social consolidation, it became more dependent on it. For larger groups ($G \geq 200$), social integration requires very high levels of social consolidation. Figure 7 shows how increasing consolidation affects the success of integration for different group sizes ($50 \leq G \leq 400$). Similarly to figure 5, in figure 7 homophily is held constant at $\alpha = 2$, and consolidation increases from $\beta = -1$ to $\beta = 3$. The solid line ($G = 50$) shows that diffusion succeeds in the range $0.5 < \beta < 1.5$. The dashed line in figure 7 ($G = 100$) corresponds to figure 5, for which

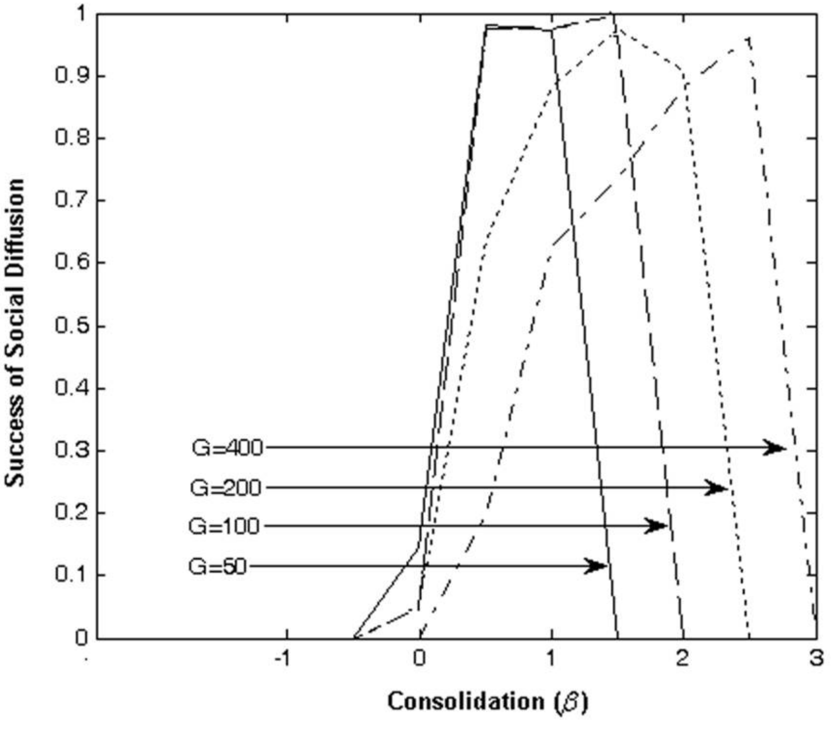


FIG. 7.—Effects of consolidation and group size on success ($H = 32$, $D = 10$, $Z = 10$, $\alpha = 2$, averaged over 100 realizations). Fraction of the population that adopts the collective behavior is shown along the $y$-axis, for increasing levels of consolidation, from random to perfect, along the $x$-axis. As group size increases, there is a greater dependence of social diffusion on consolidation. The solid line shows groups of $G = 50$, the dashed line shows groups of $G = 100$, the dotted line shows groups of $G = 200$, and the dotted-and-dashed line shows groups of $G = 400$. For the largest groups, almost perfect consolidation is needed to create networks that can support social diffusion.

diffusion has a trajectory similar to that of the solid line, but the range of success is extended up to $0.5 < \beta < 2$. For $G = 200$ (dotted line), the range of greatest success is even higher, at $0.5 < \beta < 2.5$, and for $G = 400$ (dotted-and-dashed line), successful diffusion peaks at nearly perfect consolidation, $\beta = 2.5$.

This increased dependence on social consolidation is due to the fact that when group sizes are so large, even very high levels of homophily do not guarantee the formation of social clusters and wide bridges. Even when consolidation levels are moderate ($\beta = 1$), the number of "similar" others is so large that it is very unlikely that ties will form between friends of friends to create overlapping network clusters. Increasing the level of consolidation restricts people to interacting with the same group of persons across multiple dimensions and thereby increases the likelihood that transitive ties will form. At the extreme, when group size $G = 400$, consolidation needs to be nearly perfect in order to counterbalance the effects of larger groups.

As an additional robustness test, I also examined whether the effects of social consolidation on network formation and diffusion could be independently produced by introducing a status bias into the tie formation dynamics. To do this, I altered the model to eliminate consolidation from the process of tie formation. In each dimension, I divided the population into high-status and low-status groups, based on individuals' social positions. Individuals in the top half of the dimension were considered high-status, and individuals in the bottom half were considered low-status. I then assigned each individual a probability of making ties based either on the level of homophily, as described above, or on a strong bias for attaching to members at the "high-status" end of each dimension. I systematically varied the fraction of ties that individuals formed on the basis of status bias, exploring the full range of possible values, from 0% (baseline model in which all ties are based solely on homophily) to 100% (every individual selects ties that are linked to members of the high-status group).

To provide a strong test of the effectiveness of status bias in tie formation, I did not reduce the bias if actors already had ties to high-status group members (e.g., if they were high-status actors who had already made homophilous ties or if they were low-status actors who had made or received "random" connections to high-status actors). The percentage of biased ties thus represents the lowest fraction of ties in the population that link to high-status actors.

I tested the effects of status-based tie formation across the full range of homophily values (homophilous, $\alpha = 2.5$, random, $\alpha = -1$, and heterophilous, $\alpha = -4.5$, ties). Figure 8 shows that each level of homophily presents qualitatively similar results, showing only marginal effects of status-based tie formation on social diffusion. Even with 100% status-based tie formation, diffusion never reaches more than 10%. The reason for this is

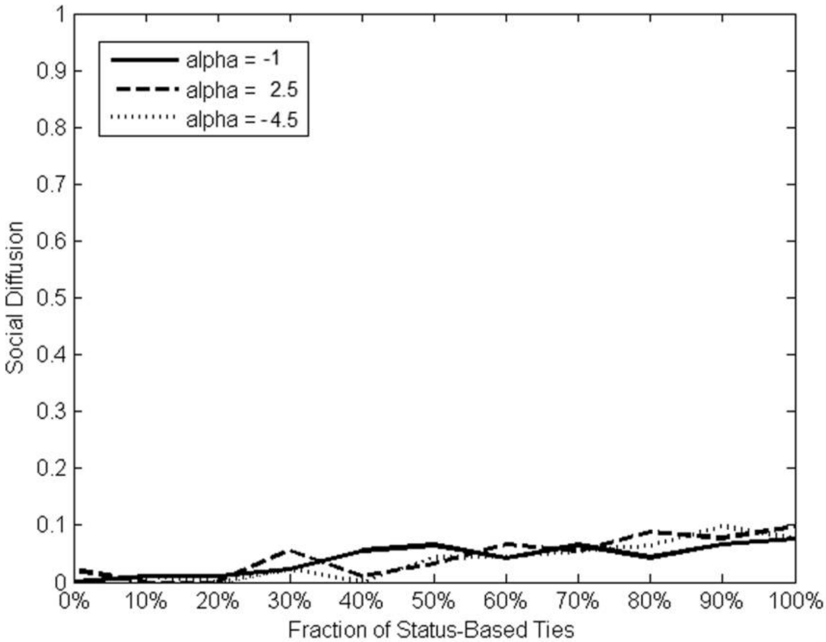


FIG. 8.—Effects of status-based tie formation on social diffusion ($H = 32$, $D = 10$, $Z = 10$, $\beta = -1$, averaged over 100 realizations). The $y$-axis shows the total fraction of the population that adopts the collective behavior, for increasing levels of status bias in tie formation, from 0% (baseline homophily model) to 100% (all links tied to high-status actors), along the $x$-axis. The solid line indicates diffusion under zero homophily ($\alpha = -1$), the dashed line shows diffusion with strong homophily ($\alpha = 2.5$), and the dotted line shows diffusion with strong heterophily ($\alpha = -4.5$). Across the full range of homophily values, there is a slight increase in adoption as status-based ties increase to 100%; however, social diffusion never reaches above 10%.

that without consolidation, individuals occupy both high and low social positions across different dimensions. Within a given dimension, strong status bias gives high-status individuals more ties and more clustering within their local groups, allowing a minimal degree of diffusion if the initial seed occurs within this cluster. Yet, in other dimensions, unrelated individuals receive the benefits of being high-status. In the emergent network, this produces moderately clustered cliques connected by large numbers of random ties. Because there is no consolidation of social positions across dimensions, there is no consistent social transitivity across people's network neighborhoods. Without overlapping ties across neighborhoods, coherent patterns of overlapping of social groups do not emerge. And in the absence of structural support from brokering structures and wide bridges, there are no pathways for social integration to spread across the population.

To sum up the results of the study, I found that for the extreme values of homophily and consolidation, one of two conditions holds: either (1) there is insufficient social clustering to support social diffusion or (2) networks form into nearly isolated clusters, preventing social diffusion. In the first case, low levels of homophily or consolidation ($\alpha < 0.5$, $\beta < 0.5$) eliminate the necessary constraints on social tie formation that allow group structures to form. This produces a highly diffuse social network akin to Durkheim's ([1897] 2005, 1997) description of a state of universal anomie, in which people have lots of social contacts, but every tie is a "weak tie." Without sufficient social structure to bind people together in cohesive groups, there is no support for the spread of shared cultural norms and practices. This inhibits the emergence of collective behaviors, normative institutions, and social integration (Durkheim 2005; Centola 2010). In the second case, high levels of both homophily and consolidation ($\alpha > 2$, $\beta > 2$) cause the social network to break apart into highly clustered groups without overlapping memberships, creating socially distinct islands across which people cannot influence one another's cultural or normative practices. This situation resembles a world in which small tribes with tightly knit internal structures are able to keep close watch over the behavior of their members. Although there may be the occasional interaction, or casual exchange, they do not share cultural or normative influences with outside groups (Barth 1969; Axelrod 1997). While these boundaries can help to maintain strong in-group cultural norms, they prevent social integration across the larger population.

The exception to these two outcomes is the middle range, where either high $\alpha$ is combined with moderate to high values of consolidation ($0.5 < \beta < 2$) or high $\beta$ is combined with low to moderate values of homophily ($0 < \alpha < 1$). When homophily and consolidation achieve this balance, the social network undergoes a remarkable shift from a topology incapable of supporting the diffusion of shared behavior across even a modest fraction of the population to one that allows behaviors to achieve rapid integration across the entire network.

## SOCIAL CONSOLIDATION, WEAK TIES, AND COMPLEX CONTAGIONS

The central implication of Blau and Schwartz's (1984) study—that reduced consolidation improves structural integration—resonates with an enormous networks literature that grows out of Granovetter's (1973) "strength of weak ties" study. From a structural point of view, the "strength" of weak ties comes from the fact that they act as bridges between distant regions of the social space, connecting people who would otherwise have no social contacts in common. That is, weak ties tend to be *long ties* (Centola and Macy 2007), which accelerate the spread of new ideas and behaviors by allowing them to "jump" across the social space. The implications of long ties for rapid diffusion were made even more salient by Watts and Strogatz's (1998) small-world model, which showed that introducing a modest fraction of long ties into a clustered social network can dramatically reduce the "degrees of separation" of the population and thereby increase the speed at which contagions spread across the network (Centola and Macy 2007).

Consistent with the weak tie and small-world models, a natural implication of Blau and Schwartz's (1984) findings is that greater numbers of crosscutting ties, and more diverse social networks, create ever-faster pathways for diffusion and integration. The reason is that the social network structures that emerge from populations with low levels of consolidation are ones with a large fraction of long ties. They are small worlds, with low characteristic path lengths, in which information and ideas can rapidly spread between people, promoting exchange across very large and very diverse populations. Highly intersecting social networks have very little clustering and therefore almost no redundancy in signaling. This abundance of long, crosscutting ties ensures that new ideas diffuse very quickly, which, in turn, maximizes the expected efficiency of social coordination, producing a highly intersecting, highly integrated population.

Yet, our findings show that although homophily and consolidation increase the level of clustering within social networks, they can also create topologies that improve the dynamics of social diffusion. The reason is that the diffusion of shared norms and practices depends on a richly overlapping social infrastructure—clustered neighborhoods and wide bridges—in order to be successful.

Long, narrow ties clearly facilitate the spread of simple contagions because it takes only a single contact with a new disease or novel idea to transmit the contagion from one region of the social space to another. However, when transmission requires social reinforcement from multiple sources, the advantage of long ties (their length) also becomes a weakness (their narrowness). The power of long ties lies in the fact that there are no alternative short paths. By contrast, clustered network ties form wide, short bridges that allow multiple signals to travel from one area to another across many, equally short pathways. These wide bridges produce overlapping patterns of social reinforcement (Centola and Macy 2007), which are not found in weak tie networks.

The more interdependent the social behavior—requiring greater support from social contacts—the more the dynamics of social diffusion depend on having wide bridges between groups. Without wide bridges to provide social reinforcement, complex social contagions can fizzle out and disappear before they are able to spread. As Centola and Macy (2007, p. 723) put it, "for simple contagions, too much clustering means too few long ties, which slows down cascades. For complex contagions, too little clustering means too few wide bridges, which not only slows down cascades but can prevent them entirely."

While our results are consistent with Centola and Macy's (2007) findings that increased clustering facilitates the spread of complex contagions, they also show that the structural implications of consolidation are more complex. Increasing network clustering through social consolidation can improve social diffusion—going from less than 1% diffusion in networks with low consolidation (fig. 6*A*) to greater than 95% diffusion in networks with moderate consolidation (fig. 6*B*). However, it can also impede diffusion, as networks with higher levels of consolidation and clustering showed diffusion rates below 5% (fig. 6*C*). Our findings indicate that there is something more than network clustering affecting the success of social diffusion in complex networks—namely, the emergence of overlapping groups with wide bridges between them.

As shown in panel *B* of figure 6, the peaks of successful diffusion along both the *x*- and *y*- axes in figure 4 correspond to the emergence of wide bridges and overlapping group structures in the social topology. When homophily and consolidation are in balance, cohesive yet interlocking social groups connect the entire population in a web of overlapping social relations (Simmel 1955). This is a subtle balance between local cohesiveness and global connectedness, which is mediated by mesolevel group structures (Hedstrom, Sandell, and Stern 2000). Unlike the "connected cave man" model (Watts 1999), these networks are not composed of small groups linked together by "shortcuts" and "weak ties" (as shown in panel *C* of fig. 6). Nor does simply increasing the number of crosscutting ties increase diffusion rates (as shown in panel *A* of fig. 6; Granovetter 1973; Blau and Schwartz 1984; Watts and Strogatz 1998). Neither small worlds nor large ones are the key to social diffusion. Rather, the social patterns that emerge from my model of network formation suggest that it is not just social ties but group structures that control the dynamics of social integration. When groups have overlapping memberships, this creates wide social bridges between different regions of the social space, which establishes relational continuity between socially remote members of the population. Thus, people are not just connected in a social network; they are embedded in a web of group affiliations (Simmel 1955), which allows social practices to diffuse from group to group and across the population.

These findings may inform recent work in organizational studies on the role of networks in promoting exchange within and across firms.[23] Increased heterophily and more weak ties have been shown to increase the rate of informational diffusion and thereby the dynamics of knowledge

[23] I am indebted to an anonymous reviewer for suggesting these implications.

exchange across organizations (Hansen 1999). However, my results suggest that these dynamics of knowledge transmission may not generalize to the spread of shared organizational values or the rapid learning of unfamiliar business practices. Rather, nontrivial levels of homophily and consolidation may be required to promote the formation and maintenance of informal overlapping group structures that can broker the spread of new normative behaviors across traditional organizational boundaries. The more complex the organization, the more important the role that may be played by these emergent network structures. Without social institutions that support the formation of overlapping groups and wide bridges, cross-group ties may succeed at fostering informational transmission but fail at promoting enduring "cultural" exchange. These results reinforce recent empirical findings that merely being connected in a social network is not enough to guarantee that behaviors will spread across groups (Centola 2010, 2011). The spread of complex behaviors depends on overlapping group structures that provide relational continuity between remote areas of the social space. When they are successful, these structures achieve a balance between supporting behavioral compliance within groups and providing conduits for behavioral transmission between them.

## CONCLUSION

Social networks emerge from social contexts: In order for people to form social ties, they must have an opportunity to interact with one another. This dependence has been appreciated by scholars in fields as varied as immigration (Kalmijn and Flap 2001), social mobility and inequality (Blau and Schwartz 1984), and epidemiology (Watts 2003). Following Blau and Schwartz (1984), the findings presented here show that homophily and consolidation do indeed put strong constraints on the contexts that people share and the kinds of social ties that form, which give rise to differential patterns of social relations—including weak ties, clustered neighborhoods, and wide bridges.

The goal of this article is to explore the implications of these emergent topological features of social networks for the collective dynamics of social diffusion. In particular, I investigated how the mechanisms that control the formation of social networks affect a population's capacity to achieve widespread social integration on complex social behaviors, such as cultural norms, shared beliefs, and collective action.

My findings show that the diffusion of shared customs, ideas, and behaviors depends on the relationship between homophily and consolidation. While homophily can range from low to high, depending on the level of social consolidation, I find that moderate to high levels of consolidation are required in order to ensure that the social structure has sufficient integrity to support the dynamics of behavioral diffusion. By contrast, the ideal of a highly intersecting society—with ubiquitous crosscutting ties, highly diverse patterns of social relations, and no consolidation—eliminates the social infrastructure that is necessary to support the diffusion of norms, practices, and beliefs across a large society.

These findings resonate with Blau and Schwartz's (1984, p. 15) "central theorem" that "many intersecting social differences promote intergroup relations." Indeed, the overlapping group structures and wide bridges that we find to be so important for social diffusion necessarily entail that people have some intersecting differences. Consistent with Blau and Schwartz's empirical findings, the model also supports the idea that interracial marriages are possible only when connections can form across traditionally consolidated groups. However, my analysis goes beyond Blau and Schwartz's thesis that reducing consolidation and weakening in-group relations improve social integration.

I find that moderate to high levels of consolidation are needed in order to create the overlapping group structures that facilitate widespread social diffusion. On the basis of my results, a reinterpretation of Blau and Schwartz's (1984) empirical findings might suggest that while some crosscutting ties are necessary to promote interaction across groups, wider normative acceptance of intermarriage may depend on bridges between groups that are sufficiently wide to permit reinforcing influences both within and across communities in support of these relationships. For instance, poorly integrated long ties across a network are unlikely to affect normative acceptance of new marriage practices. However, overlapping clusters of social contacts across traditionally isolated groups can create patterns of reinforcement that may successfully foster collective acceptance of intermarriage.

These results may also have implications for the growth of social movements and the spread of social consensus on issues such as civil rights, environmental protection, and gender discrimination (McPherson and Smith-Lovin 1987; Ridgeway and Balkwell 1997). On the one hand, initiating a social movement often requires organizing a critical mass of participants through coordinated consensus among interested parties (Marwell and Oliver 1993). On the other hand, once a critical mass forms, the subsequent growth of participation depends on belief in the movement spreading across diverse communities in a social network (Centola and Macy 2007). My findings suggest that too little consolidation may forestall the initiation of critical mass collective action by preventing the creation of coherent groups with coordinated interests capable of establishing a movement's initial foothold into a population. Correspondingly, my results also suggest that too much consolidation can limit the growth of a movement by restricting the initially interested groups to segregated regions of the social network, preventing them from mobilizing broad support for their cause. The widespread growth of interest and participation in social movements may thus depend on moderate levels of consolidation and homophily, which create overlapping groups and wide bridges throughout a society, facilitating both the initial emergence of a critical mass and the subsequent spread of social reinforcement for the new beliefs across the population.

Methodologically, my goal of combining Blau's classical approach to social structure with a more dynamical approach to understanding social integration as a network diffusion process may provide some value for scholars looking to connect rich structural data with implications for social diffusion. The empirical problem of accurately measuring social networks has endured for decades and has become more poignant as network studies have increasingly focused on large populations with more complex topologies (Watts and Strogatz 1998; Albert et al. 1999; Centola and Macy 2007). While getting accurate measurements of social network topologies remains a difficult problem, Blau and Schwartz (1984) show that homophily and consolidation are readily measurable structural features of large, complex populations. My results on the formal implications of these structural factors for the formation of networks and the social processes that take place across them may suggest a means of connecting these measurable properties of large populations with the collective social dynamics that they exhibit.

In conclusion, I appreciate that the recent growth of research on social networks has eclipsed much of the related, and very interesting, work on how other kinds of social structures affect collective outcomes. While social networks provide a powerful lens into social structure, I hope that this study will help to connect the growing literature on network-based approaches to studying social diffusion with other structural approaches, such as institutional and demographic research, aimed at understanding the complex effects of social structure on the dynamics of collective behavior.

## APPENDIX

### The Formal Model

The model of social network formation is based on the premise that people in social networks not only have social ties but also have social identities, which define their proximity or distance from others within a dimension of social life. This approach extends the model of network formation developed by Watts et al. (2002) in two ways. First, it introduces consolidation into the model, which allows us to explore the range of network structures generated by the interaction between homophily and consolidation. While the Watts et al. model focuses exclusively on measuring the length of completed messages (see Travers and Milgram 1969), I study the topological structure of the networks that emerge. Second, I also introduce the Centola and Macy (2007) complex contagion model into the architecture, which allows us to identify the effects of the emergent topologies on social diffusion. Consistent with Watts et al. (2002, p. 1303), the approach used here follows from six contentions about social networks.

*Contention 1*

Individuals' social identities are defined by their association with, and participation in, social groups.

*Contention 2*

Each dimension of social life can be partitioned into groups using a hierarchical representation, as shown in figure 1. This representation of a social dimension is not the actual network but is a cognitive construct for measuring social distance between individuals (Watts et al. 2002, p. 1303). As in figure 1, this partitioning ends with specialized subgroups that are relatively small and socially proximate. The parameter $G$ refers to the number of people in each subgroup. The social distance between individuals $i$ and $j$, $x_{ij}$, within a dimension of social life is defined as their closest partition level: $x_{ij} = 1$ if $i$ and $j$ belong to the same group, $x_{ij} = 2$ if $i$ and $j$ are both under the next-highest partition, and so forth. As shown in figure 1, a hierarchy is fully defined by the number of levels of partitioning $L$ and the branching ratio $B$. The number of groups, or social positions, in a dimension is given by $B^{L-1}$.

*Contention 3*

The actual social network is created on the basis of the principle that group memberships are the primary basis for social interaction and therefore the formation of social ties. The probability that a social tie will form between individuals $i$ and $j$ increases with their social proximity. This is modeled by choosing an individual $i$ at random and a distance $x$ with probability $p(x) = ce^{-\alpha x}$, where $\alpha$ is a tunable parameter that controls homophily, and $c$ is a normalizing constant. A node $j$ is then chosen randomly from among all nodes at distance $x$ from $i$. This process is repeated until a network is constructed in which all individuals have an average number of friends, $Z$. When homophily is high, $e^{-\alpha} \ll 1$, individuals will be connected only to people within their immediate subgroup ($x_{ij} = 1$ for all $i$ and $j$ who have network ties to one another). By contrast, when there is no homophily, $e^{-\alpha} = B$, ties are equally likely to form at all distances.

*Contention 4*

There are multiple dimensions of social life. The parameter $D \geq 1$ determines the number of dimensions. As shown in figure 2, each dimension is represented as its own independent hierarchy. Each individual is randomly assigned to a position in dimension $h_1$. The correlation between an individual's social positions across dimensions (i.e., social consolidation) is modeled by assigning an individual's social positions in $h_2 - h_H$ at distance $y$ from her position in $h_1$ with the probability $p(y) = ce^{-\beta y}$, where $\beta$ is a tunable social consolidation parameter, and $c$ is a normalizing constant. When social consolidation is very high, $e^{-\beta} \ll 1$, individuals will be located in the same subgroup, or social position, across all dimensions (i.e., $y^{h_1 h_2} = y^{h_1 h_3} = \cdots = y^{h_1 h_H} = 1$). By contrast, when there is no consolidation, $e^{-\beta} = B$, people are equally likely to be located in any social position, making social positions uncorrelated across dimensions.

*Contention 5*

Social distance, unlike Euclidean space, is not transitive. Two people $i$ and $j$ may have high social similarity in dimension $h_1$, while $i$ may be close to $k$ in dimension $h_2$. But, this does not imply that $j$ and $k$ are close to each other in any dimension. The ultrametric distance used here to measure the social distance between people preserves this intransitivity of social relations (Watts et al. 2002).

*Contention 6*

People have thresholds for adopting new behaviors (Granovetter 1978; Watts 2002, 2003; Centola and Macy 2007). An individual's threshold, $T = r/z$, is represented as the number of contacts who are adopters, $r$, over the total number of contacts, $z$. This notation distinguishes between 1/8 and 6/48, which represent the same proportion of adopters; but the former corresponds to a threshold for a simple contagion while the latter corresponds to a complex contagion (Centola and Macy 2007). The dynamics of social integration are modeled as a social diffusion process, where individuals decide whether or not to adopt shared practices and behaviors. An individual's adoption decision is contingent on having a sufficient number of her social contacts adopt the behavior that her threshold of adoption, $T$, is triggered.

Since the principal goal of this study is to understand how the social forces of homophily and consolidation affect the dynamics of social diffusion, I begin by fixing the model parameters that are not related to homophily or consolidation. These include the number of dimensions of social life ($D$), the heterogeneity of social positions (i.e., the number of groups in each dimension, $B^{L-1}$), and the size of the groups ($G$). Following the initial analysis, I also explore the interactions of these parameters with homophily and consolidation and highlight any deviations from the initial results. The parameters for homophily ($\alpha$) and consolidation ($\beta$) are systematically explored within the range $\alpha = [-1, 3]$ (ranging from random tie formation to nearly perfect homophily in tie formation) and $\beta = [-1, 3]$ (ranging from no correlation in social positions to nearly perfect correlation of positions across dimensions). Random assignment to social positions (zero correlation across dimensions) occurs when $\beta \approx -1$, and random selection of social ties (zero homophily in tie formation) occurs when $\alpha \approx -1$.

## REFERENCES


Albert, R., H. Jeong, and A. Barabasi. 1999. "The Diameter of the World-Wide-Web." *Nature* 401:130–31.

Axelrod, Robert. 1984. *The Evolution of Cooperation*. New York: Basic Books.

———. 1997. "The Dissemination of Culture: A Model with Local Convergence and Global Polarization." *Journal of Conflict Resolution* 41:203–26.

Backstrom, Lars, Dan Huttenlocher, Xiangyang Lan, and Jon Kleinberg. 2006. "Group Formation in Large Social Networks: Membership, Growth, and Evolution." Pp. 44–54 in *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York: ACM.

Barth, Fredrik. 1969. *Ethnic and Group Boundaries*. Long Grove, Ill.: Waveland.

Blau, P. M. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. New York: Free Press.

Blau, P. M., and J. E. Schwartz. 1984. *Crossing Social Circles*. Orlando, Fla.: Academic Press.

Blum, T. C. 1985. "Structural Constraints on Interpersonal Relations: A Test of Blau's Macrosociological Theory." *American Journal of Sociology* 91:511–21.

Boorman, S. 1974. "Island Models for Takeover by a Social Trait Facing a Frequency-Dependent Selection Barrier in a Mendelian Population." *Proceedings of the National Academy of Sciences USA* 71:2103–7.

Boyd, R., and P. J. Richerson. 1985. *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.

Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, Mass.: Harvard University Press.

Castro, L., and M. A. Toro. 2004. "Culture as an Evolutionary Process: From Primate Social Learning to Human Culture." *Proceedings of the National Academy of Sciences* 101:10235–40.

Centola, Damon. 2010. "The Spread of Behavior in an Online Social Network Experiment." *Science* 329:1194–97.

———. 2011. "An Experimental Study of Homophily in the Adoption of Health Behavior." *Science* 334:1269–72.

Centola, Damon, Victor Eguiluz, and Michael Macy. 2007. "Cascade Dynamics of Complex Propagation." *Physica A* 374:449–56.
Centola, Damon, Juan Carlos Gonzalez-Avella, Victor Eguiluz, and Maxi San Miguel. 2007. "Homophily, Cultural Drift, and the Co-evolution of Cultural Groups." *Journal of Conflict Resolution* 51:905–29.
Centola, Damon, and Michael Macy. 2007. "Complex Contagions and the Weakness of Long Ties." *American Journal of Sociology* 113:702–34.
Centola, Damon, and Arnout van de Rijt. 2015. "Choosing Your Network: Social Preferences in an Online Health Community." *Social Science and Medicine* 125:19–31.
Centola, Damon, Robb Willer, and Michael Macy. 2005. "The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms." *American Journal of Sociology* 110: 1009–40.
Collins, Randall. 1993. "Emotional Energy as the Common Denominator of Rational Action." *Rationality and Society* 5:203–30.
Currarini, Sergio, Matthew O. Jackson, and Paolo Pin. 2010. "Identifying the Roles of Race-Based Homophily in High School Friendship Network Formation." *Proceedings of the National Academy of Sciences* 107:4857–61.
Dodds, P. S., R. Muhamad, and Duncan J. Watts. 2003. "An Experimental Study of Search in Global Social Networks." *Science* 301:827–29.
Durkheim, E. (1893) 1997. *The Division of Labor in Society*, translated by Lewis A. Coser. New York: Free Press.
———. (1897) 2005. *Suicide: A Study in Sociology*, translated by J. A. Spaulding and G. Simpson. New York: Routledge.
Feld, Scott L. 1981. "The Focused Organization of Social Ties." *American Journal of Sociology* 86:1015–35.
Friedkin, Noah E. 1998. *A Structural Theory of Social Influence*. New York: Cambridge University Press.
———. 2001. "Norm Formation in Social Influence Networks." *Social Networks* 23 (3): 167–89.
Granovetter, M. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78: 1360–80.
———. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83:1420–43.
Grindereng, M. P. 1967. "Fashion Diffusion." *Journal of Home Economics* 59:171–74.
Hansen, Morten T. 1999. "The Search-Transfer Problem: The Role of Weak Ties in Sharing Knowledge across Organization Subunits." *Administrative Science Quarterly* 44:82–111.
Hedström, Peter, Rickard Sandell, and Charlotta Stern. 2000. "Mesolevel Networks and the Diffusion of Social Movements: The Case of the Swedish Social Democratic Party." *American Journal of Sociology* 106:145–72.
Kalmijn, Matthijs, and Henk Flap. 2001. "Assortative Meeting and Mating: Unintended Consequences of Organized Settings for Partner Choices." *Social Forces* 79:1289–1312.
Keeling, M. J. 1999. "The Effects of Local Spatial Structure on Epidemiological Invasions." *Proceedings of the Royal Society of London, B*, 266:859–67.
Kim, Hyojoung, and Peter Bearman. 1997. "The Structure and Dynamics of Movement Participation." *American Sociological Review* 62:70–93.
Klandermans, Bert. 1988. "Union Action and the Free-Rider Dilemma." Pp. 77–92 in *Research in Social Movements, Conflict and Change*. Vol. 10, *Social Movements as a Factor of Change in the Contemporary World*, edited by L. Kriesberg and B. Misztal. Greenwich, Conn.: JAI Press.
Kleinberg, Jon. 2000. "Navigation in a Small World." *Nature* 406:845.

Klemm, Konstantin, Victor M. Eguíluz, Raúl Toral, and Maxi San Miguel. 2003. "Global Culture: A Noise Induced Transition in Finite Systems." *Physical Review E* 67: 045101 (R).

Knottnerus, J. David, and Jian Guan. 1997. "The Works of Peter M. Blau: Analytical Strategies, Developments and Assumptions." *Sociological Perspectives* 40:109–28.

Kossinets, Gueori, and Duncan J. Watts. 2009. "Origins of Homophily in an Evolving Social Network." *American Journal of Sociology* 115:405–50.

Liljeros, F., C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg. 2001. "The Web of Human Sexual Contacts." *Nature* 411:907–8.

Macy, Michael W. 1990. "Learning Theory and the Logic of Critical Mass." *American Sociological Review* 55:809–26.

Marwell, Gerald, and Pamela Oliver. 1993. *The Critical Mass in Collective Action: A Micro-Social Theory*. New York: Cambridge University Press.

McPherson, J. Miller. 2004. "A Blau Space Primer: Prolegomenon to an Ecology of Affiliation." *Industrial and Corporate Change* 13:263–80.

McPherson, J. Miller, Pamela A. Popielarz, and Sonja Drobnic. 1992. "Social Networks and Organizational Dynamics." *American Sociological Review* 57:153–70.

McPherson, J. Miller, and Lynn Smith-Lovin. 1987. "Homophily in Voluntary Organizations: Status Distance and the Composition of Face-to-Face Groups." *American Sociological Review* 52:370–79.

McPherson, J. Miller, Lynn Smith-Lovin, and James Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.

Moody, James, and Douglas R. White. 2003. "Structural Cohesion and Embeddedness: A Hierarchical Conception of Social Groups." *American Sociological Review* 68:103–27.

Newman, M. E. 2000. "Models of the Small World: A Review." *Journal of Statistical Physics* 101:819–41.

Newman, Mark, A. L. Barabasi, and Duncan Watts. 2006. *The Structure and Dynamics of Networks*. Princeton, N.J.: Princeton University Press.

Newman, Mark E., and Duncan Watts. 1999. "Scaling and Percolation in the Small-World Network Model." *Physical Review E* 60:7332–42.

Popielarz, Pamela, and J. Miller McPherson. 1995. "On the Edge or in Between: Niche Position, Niche Overlap, and the Duration of Voluntary Memberships." *American Journal of Sociology* 101:698–720.

Reagans, Ray. 2005. "Preferences, Identity, and Competition: Predicting Tie Strength from Demographic Data." *Management Science* 51:1374–83.

Reagans, Ray, Ezra Zuckerman, and Billy McEvily. 2004. "How to Make the Team? Social Networks vs. Demography as Criteria for Designing Effective Teams." *Administrative Science Quarterly* 49:101–33.

Ridgeway, Cecilia, and J. Balkwell. 1997. "Group Processes and the Diffusion of Status Beliefs." *Social Psychological Quarterly* 60:14–31.

Rogers, Everett M. 1995. *Diffusion of Innovations*. New York: Free Press.

Sampson, Robert J. 1984. "Group Size, Heterogeneity, and Intergroup Conflict: A Test of Blau's Inequality and Heterogeneity." *Social Forces* 62:618–39.

Schelling, T. 1978. *Micromotives and Macrobehavior*. New York: Norton.

Shepard, Roger N. 1987. "Toward a Universal Law of Generalization for Psychological Science." *Science* 237:1317–23.

Shrum, Wesley, Neil H. Cheek, Jr., and Saundra M. Hunter. 1988. "Friendship in School: Gender and Racial Homophily." *Sociology of Education* 61:227–39.

Simmel, Georg. (1922) 1955. *Conflict and the Web of Group Affiliations*, translated and edited by Kurt Wolff. Glencoe, Ill.: Free Press.

———. 1950. *The Sociology of Georg Simmel*, compiled and translated by Kurt Wolff. Glencoe, Ill.: Free Press.

Skvoretz, John. 1983. "Salience, Heterogeneity, and Consolidation of Parameters: Civilizing Blau's Primitive Theory." *American Journal of Sociology* 48:360–75.

Tenenbaum, J. B., and T. L. Griffiths. 2002. "Generalization, Similarity, and Bayesian Inference." *Behavioral and Brain Sciences* 24:629–40.

Travers, Jeffrey, and Stanley Milgram. 1969. "An Experimental Study of the Small World Problem." *Sociometry* 32 (4): 425–43.

Verbrugge, Lois M. 1977. "The Structure of Adult Friendship Choices." *Social Forces* 62:78–83.

Watts, Duncan J. 1999. *Small Worlds: The Dynamics of Networks between Order and Randomness*. Princeton, N.J.: Princeton University Press.

———. 2002. "A Simple Model of Global Cascades on Random Networks." *Proceedings of the National Academy of Sciences* 99:5766–71.

———. 2003. *Six Degrees: The Science of a Connected Age*. New York: Norton.

Watts, Duncan J., Peter Sheridan Dodds, and M. E. J. Newman. 2002. "Identity and Search in Social Networks." *Science* 296:1302–5.

Watts, Duncan J., and Steven H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393:440–42.