

# Influentials, Networks, and Public Opinion Formation

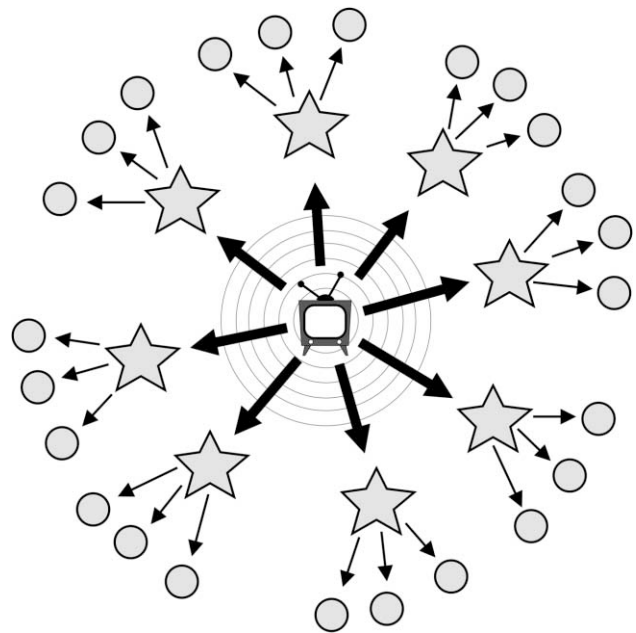
DUNCAN J. WATTS  
PETER SHERIDAN DODDS\*

A central idea in marketing and diffusion research is that influentials—a minority of individuals who influence an exceptional number of their peers—are important to the formation of public opinion. Here we examine this idea, which we call the “influentials hypothesis,” using a series of computer simulations of interpersonal influence processes. Under most conditions that we consider, we find that large cascades of influence are driven not by influentials but by a critical mass of easily influenced individuals. Although our results do not exclude the possibility that influentials can be important, they suggest that the influentials hypothesis requires more careful specification and testing than it has received.

In the 1940s and 1950s, Paul Lazarsfeld, Elihu Katz, and colleagues (Katz and Lazarsfeld 1955; Lazarsfeld, Berelson, and Gaudet 1968) formulated a breakthrough theory of public opinion formation that sought to reconcile the role of media influence with the growing realization that, in a variety of decision-making scenarios, ranging from political to personal, individuals may be influenced more by exposure to each other than to the media. According to their theory, illustrated schematically in figure 1, a small minority of “opinion leaders” (stars) act as intermediaries between the mass media and the majority of society (circles). Because information, and thereby influence “flows” from the media through opinion leaders to their respective followers, Katz and Lazarsfeld (1955) called their model the “two-step flow” of communication, in contrast with the then paradigmatic one-step, or “hypodermic,” model that treated individuals as atomized objects of media influence (Bineham 1988).

In the decades after the introduction of the two-step flow, the idea of opinion leaders, or “influentials” as they are also called (Merton 1968), came to occupy a central place in the literatures of the diffusion of innovations (Coleman, Katz, and Menzel 1966; Rogers 1995; Valente 1995), communications research (Weimann 1994), and marketing (Chan and

FIGURE 1  
SCHEMATIC OF THE TWO-STEP FLOW MODEL  
OF INFLUENCE



Misra 1990; Coulter, Feick, and Price 2002; Myers and Robertson 1972; Van den Bulte and Joshi 2007; Vernet 2004). By the late 1960s, the theory had been hailed as one of the most important formulations in the behavioral sciences (Arndt 1967), and by the late 1970s, according to Gitlin (1978), the two-step flow had become the “dominant paradigm” of media sociology. According to Weimann (1994),

\*Duncan J. Watts is professor of sociology, Columbia University, New York, NY 10025 (djw24@columbia.edu). Peter Sheridan Dodds is assistant professor of mathematics and statistics, University of Vermont, Burlington VT 05404 (peter.dodds@uvm.edu). The authors acknowledge the helpful input of the editor, associate editor, and reviewers. This research was supported in part by the National Science Foundation (SES-0094162 and SES-0339023), and the McDonnell Foundation.

John Deighton served as editor and Tulin Erdem served as associate editor for this article.

Electronically published May 30, 2007

over 3,900 studies of influentials, opinion leaders, and personal influence were conducted in the decades after Katz and Lazarsfeld's seminal Decatur study, leading Burt (1999, 38) to comment that the two-step flow had become "a guiding theme for diffusion and marketing research." More recently still, Roch (2005, 110) has concluded that "in business and marketing, the idea that a small group of influential opinion leaders may accelerate or block the adoption of a product is central to a large number of studies."

But what exactly does the two-step flow say about influentials, and how precisely do they exert influence over the (presumably much larger) population of noninfluentials? In the remainder of this article we argue that, although the dual concepts of personal influence and opinion leadership have been extensively documented, it is nevertheless unclear exactly how, or even if, the influentials of the two-step flow are responsible for diffusion processes, technology adoption, or other processes of social change. By simulating a series of formal models of diffusion that are grounded in the mathematical social science literature, we find that there are indeed conditions under which influentials are likely to be disproportionately responsible for triggering large-scale "cascades" of influence and that, under these conditions, the usual intuition regarding the importance of influentials is supported. These conditions, however, appear to be the exception rather than the rule—under most conditions that we consider, influentials are only modestly more important than average individuals. In the models that we have studied, in fact, it is generally the case that most social change is driven not by influentials but by easily influenced individuals influencing other easily influenced individuals. Based on these results, we argue that, although our models are at best a simplified and partial representation of a complex reality, they nevertheless highlight that claims regarding the importance of influentials should rest on carefully specified assumptions about who influences whom and how.

## THE INFLUENTIALS HYPOTHESIS

Katz and Lazarsfeld (1955, 3) originally defined *opinion leaders* as "the individuals who were likely to influence other persons in their immediate environment," and this definition remains in use, more or less unchanged (Grewal, Mehta, and Kardes 2000, 236). It is important to note that opinion leaders are not "leaders" in the usual sense—they do not head formal organizations nor are they public figures such as newspaper columnists, critics, or media personalities, whose influence is exerted indirectly via organized media or authority structures. Rather their influence is direct and derives from their informal status as individuals who are highly informed, respected, or simply "connected." As Keller and Berry (2003, 1) specify: "It's not about the first names that come to mind when you think about the people with influence in this country—the leaders of government, the CEO's of larger corporations, or the wealthy. Rather, it's about millions of people . . . who shape the opinions and trends in our country." Oprah Winfrey, in other words, may or may not influence public opinion in various ways, but that is a question primarily of media, not

interpersonal, influence, and therefore is the subject of a different discussion.

Although the notion of opinion leadership seems clear, precisely how the influence of opinion leaders over their "immediate environment" shapes opinions and trends across entire communities, or even a country, is not specified by the two-step flow model itself. Often it is described simply as "a process of the moving of information from the media to opinion leaders, and influence moving from opinion leaders to their followers" (Burt 1999, 38), where the mechanics of the process itself are either left unspecified or, alternatively, are asserted to derive from some diffusion process. Kelly et al. (1991, 168), for example, claim that "diffusion of innovation theory posits that trends and innovations are often initiated by a relatively small segment of opinion leaders in the population," and similar statements appear in fields as diverse as marketing (Chan and Misra 1990; Coulter et al. 2002; Van den Bulte and Joshi 2007), public health (Doumit et al. 2007; Kelly et al. 1991; Moore et al. 2004, 189; Soumerai et al. 1998, 1358), and political behavior (Nisbet 2006; Roch 2005), as well as being featured in popular books (Barabasi 2002; Gladwell 2000) and commercial marketing publications (Burson-Marsteller 2001; Keller and Berry 2003; Rand 2004).

On closer inspection, however, the diffusion of innovations literature does not specify any such theory. Rogers (1995, 281) does indeed state the following: "The behavior of opinion leaders is important in determining the rate of adoption of an innovation in a system. In fact, the S-shape of the diffusion curve occurs because once opinion leaders adopt and tell others about the innovation, the number of adopters per unit time takes off." But Rogers's claim does not necessarily follow from the dynamics of diffusion processes—S-shaped diffusion curves do not, in fact, require opinion leaders at all. The "Bass model" (Bass 1969), for example, which is also popular in the marketing literature (Lehmann and Esteban-Bravo 2006), invariably generates S-shaped diffusion curves; yet it does so within an entirely homogenous population. Nor do other kinds of diffusion models (Bailey 1975; Dodds and Watts 2005; Dodson and Muller 1978; Granovetter 1978; Shrager, Hogg and Huberman 1987; Young 2006) require opinion leaders or any kind of special individuals in order to generate S-shaped diffusion curves, which are characteristic of virtually any self-limiting, contagious process.

The absence of special individuals in formal models of diffusion, of course, does not necessarily mean that they do not arise in a real diffusion process or that they do not play an important role. It does suggest, however, that the matter has not been resolved either by the two-step flow itself or by diffusion of innovations theory. To what extent, therefore, does the observation that some people are more influential than others in their immediate environment translate to the much stronger and more interesting claim that some special group of influentials plays a critical, or at least important, role in forming and directing public opinion? To address this question, we will study, using computer simulations, a

series of mathematical models of interpersonal influence that invoke a variety of assumptions about both the nature of interpersonal influence and influence networks. We emphasize that the models we consider, while plausible and somewhat general, are neither realistic nor exhaustive in their consideration of possible alternatives. However, we have consciously relaxed some of our more restrictive assumptions; thus, to the extent that different assumptions consistently generate similar conclusions, these conclusions might be expected to have reasonably broad applicability.

## A SIMPLE MODEL OF INTERPERSONAL INFLUENCE

We start by assuming that each individual  $i$  must make a decision with regard to some issue  $X$ . Following an extensive literature that encompasses sociology (Granovetter 1978), economics (Blume and Durlauf 2003; Brock and Durlauf 2001; Durlauf 2001; Schelling 1973; Young 1996), social psychology (Latané and L'Herrou 1996), political science (Kuran 1991), marketing (Goldenberg, Libai, and Muller 2001; Mayzlin 2002), and mathematical physics (Borghesi and Bouchaud forthcoming; Watts 2002), we focus on binary decisions that exhibit “positive externalities,” meaning that the probability that individual  $i$  will choose alternative B over A will increase with the relative number of others choosing B. Positive externalities arise in a wide variety of areas of interest to marketing and diffusion researchers, including “network effects” (Liebowitz and Margolis 1998), coordination games (Latané and L'Herrou 1996; Oliver and Marwell 1985), social proof (Cialdini 2001); “learning from others” (Bikhchandani, Hirshleifer, and Welch 1992; Sarganik, Dodds, and Watts 2006), and conformity pressures (Bernheim 1994; Bond and Smith 1996; Cialdini and Goldstein 2004). Furthermore, as Schelling (1978), Granovetter (1978), and others (Banerjee 1992; Bikhchandani et al. 1992; Watts 2003) have argued, binary decisions can be applied to a surprisingly wide range of real world situations. Thus, while the class of binary decisions with positive externalities is not completely general—for example, it excludes cases in which individuals must choose simultaneously between many alternatives (De Vany and Walls 1996; Hedstrom 1998) and also excludes “anticoordination” (Arthur 1994) or “snob” (Leibenstein 1950) behavior, for which negative externalities apply—it is an important and reasonably general case to consider.

### Threshold Rule

Within this class, a widely studied decision rule, and the one that we will consider initially (we will consider an alternative later), is called a “threshold rule.” This rule posits that individuals will switch from A to B only when sufficiently many others have adopted B in order for the perceived benefit of adopting a new innovation to outweigh the perceived cost (Lopez-Pintado and Watts 2007; Morris 2000; Schelling 1973). More formally, when  $b_i$ , the fraction of  $i$ 's sample population that has adopted B is less than  $i$ 's

threshold  $\phi_i$ , the probability that  $i$  will adopt B is zero, and when it exceeds  $\phi_i$ , it jumps to one; that is,

$$\Pr[\text{adopt B}] = \begin{cases} 1 & \text{if } b_i \geq \phi_i, \\ 0 & \text{if } b_i < \phi_i, \end{cases}$$

where individual differences with respect to subject matter expertise, strength of opinion, personality traits, media exposure, or perceived adoption costs may correspond to heterogeneous thresholds  $b_i$  (Valente 1995).

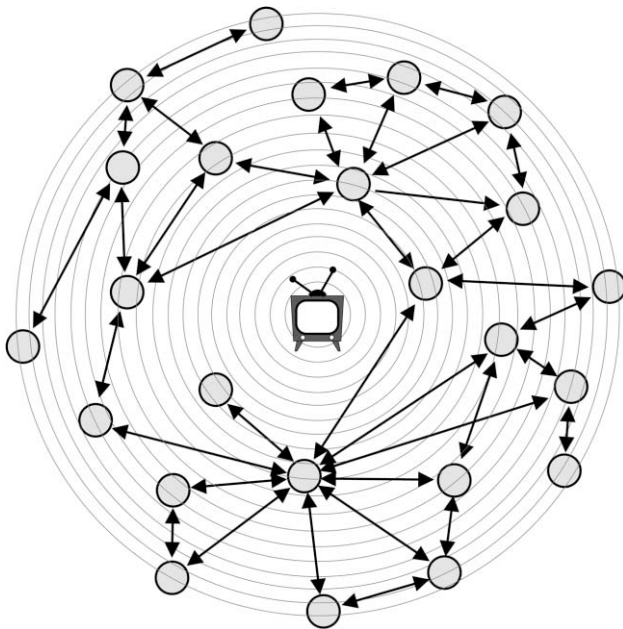
### Influence Networks

In addition to describing a rule for how individuals influence each others' decisions, we also need to specify who influences whom—that is, we require a formal description of the associated influence network. Unfortunately, empirical evidence regarding real world influence networks is limited. Although a number of sociometric studies of influence have been conducted, beginning with Coleman et al.'s (1957) study of the diffusion of tetracycline, the resulting analyses have tended to focus on individual-level (Burt 1987) rather than network properties. More recently, a considerable amount of empirical work has been dedicated to measuring the properties of large-scale networks (Newman 2003; Watts 2004); however, only a handful of studies (Godes and Mayzlin 2004; Leskovec, Adamic, and Huberman 2007) have attempted to study word-of-mouth influence empirically. These approaches, moreover, while promising, are not yet capable of measuring network structure in any detail. Finally, a small number of human subjects experiments (Kearns, Suri, and Montfort 2006; Latané and L'Herrou 1996) have shed light on the relationship between network structure and group coordination processes. However, the networks involved were artificially generated by the experimenters; thus, they do not shed light on the structure of real world influence networks.

In the absence of clear empirical evidence regarding the structure of influence networks, we assume that each individual  $i$  in a population of size  $N$  influences  $n_i$  others, chosen randomly, where  $n_i$  is drawn from an influence distribution  $p(n)$ , whose average  $n_{\text{avg}}$  is assumed to be much less than the size of the total population—that is,  $n_{\text{avg}} \ll N$ . We emphasize that, although  $n_i$  is mathematically equivalent to the notion of “acquaintance volume”  $k_i$ , familiar to network analysts (Wasserman and Faust 1994),  $n_i$  refers not to how many others node  $i$  knows but to how many others  $i$  influences with respect to the particular issue,  $X$ , at hand. Thus  $n_i$  should be thought of as some (possibly complicated) function not only of  $i$ 's acquaintance volume but also of  $i$ 's personal characteristics, subject matter expertise, authority with respect to issue  $X$ , and even the characteristics of the other individuals in  $i$ 's community (Katz and Lazarsfeld 1955; Rogers 1995).

The resulting influence network, illustrated schematically in figure 2, differs from the two-step flow schematic of figure 1 in two important ways. First, whereas in figure

**FIGURE 2**  
SCHEMATIC OF NETWORK MODEL OF INFLUENCE



1 influence can only flow from opinion leaders to followers, in figure 2, it can flow in either direction. Second, in figure 2 influence can propagate for many steps, whereas in figure 1 it can propagate only two. We note, however, that, in both cases, figure 2 is consistent with available empirical evidence—arguably more so than figure 1. Numerous studies, including that of Katz and Lazarsfeld (1955), suggest that opinion leaders and followers alike are exposed to mixtures of interpersonal and media influence (Troidahl and Dam 1965) and that differences in influence are more appropriately described on a continuum than dichotomously (Lin 1973). Furthermore, while the implications of multistep flow for the influentials hypothesis have not been studied, multistep flow itself has been recognized as a likely feature of most diffusion processes (Menzel and Katz 1955; Robinson 1976). Brown and Reingen (1987), for example, found that, even in a relatively small population, 90% of recommendation chains extended over more than one step and 38% involved at least four individuals.

Although our model therefore embodies some important qualitative features of real world interpersonal influence networks, it nevertheless contains two assumptions regarding the structure of these networks that merit critical examination in light of our objectives. First, the influence distribution  $p(n)$ , which is necessarily Poisson (Solomonoff and Rapoport 1951), exhibits relatively little variation around its average. Influentials in such a world, while clearly more influential than average, are rarely many times more influential. Second, aside from the distribution of influence, the network exhibits no other structure—it is entirely random. Although neither

of these assumptions is demonstrably incorrect, neither is clearly correct either—the empirical evidence is unfortunately inconclusive. Thus we will also present in a later section two variations of the basic model that relax both the homogeneity and the randomness assumptions.

Another advantage of formally defining an influence network, even with such a simple model, is that we can now define more precisely what we mean by an “influential.” Previous empirical work has addressed the question of who should be considered influential, but a clear answer remains elusive (Weimann 1991). Classical studies like those of Coleman et al. (1957) and Merton (1968) suggested that individuals who directly influence more than three or four of their peers should be considered influentials, while recent market research studies have concluded that the number may be as high as 14 (Burson-Marsteller 2001). Other studies, by contrast, define influentials in purely relative terms: Keller and Berry (2003), for example, define influentials as scoring in the top 10% of an opinion leadership test, while Coulter et al. (2002), using a similar test, treat the top 32% as influentials.

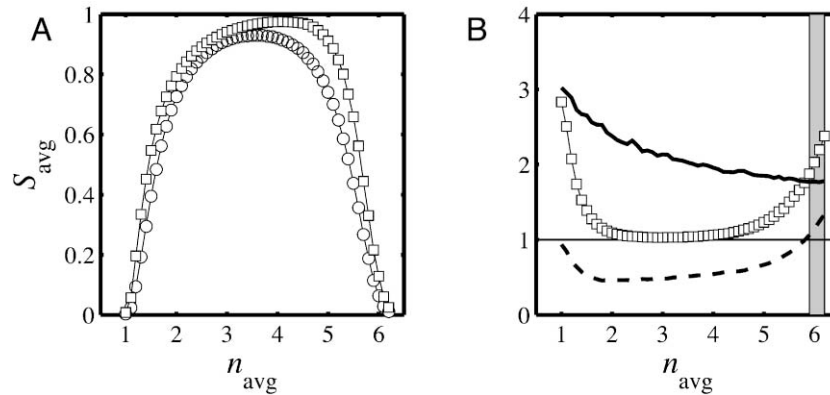
Here we follow the latter approach and define an influential as an individual in the top  $q\%$  of the influence distribution  $p(n)$ . From a theoretical perspective, any particular value of  $q$  that we specify is necessarily arbitrary—indeed, we have already argued that dichotomies such as that between opinion leaders and followers are neither theoretically derived nor empirically supported. Our purpose here, however, is not to defend any particular definition of influentials but to examine the claim that influentials—defined in some reasonable, self-consistent manner—determine the outcome of diffusion processes. From this perspective, therefore, our definition has the advantage (over definitions that rely on absolute numbers) that it can be applied consistently to influence networks of different average densities  $n_{\text{avg}}$  and also different distributions  $p(n)$ . In all results presented here, we choose  $q = 10\%$ —a number that is consistent with previous studies (Keller and Berry 2003)—but we have also studied a wide range of values of  $q$  and have determined that our conclusions do not depend sensitively on the specific choice.

### Dynamics of Influence

Our model proceeds from an initial state in which all  $N$  individuals are inactive (state 0), with the exception of a single, randomly chosen *initiator*  $i$ , who is activated (state 1) exogenously. Depending on the model parameters and also on the particular (randomly chosen) properties of  $i$ 's neighbors, this initial activation may or may not trigger some additional endogenous activations. Subsequently, these newly activated neighbors may activate some of their own neighbors, who may, in turn, trigger more activations still, and so on, generating a sequence of activations, called a “cascade” (Watts 2002). When all activations associated with a single cascade have occurred, its size can be determined simply as the total cumulative number of activations. By repeating this process many times, where each time the population, the corresponding influence network, and the

FIGURE 3

RELATIVE IMPACT OF INFLUENTIALS AS INITIATORS FOR RANDOM INFLUENCE NETWORKS AND THE THRESHOLD MODEL OF INFLUENCE



NOTE.—A, expected cascade size triggered by an influential (squares) and average (circles) individual, respectively, as a function of the density of the influence network; B, the vertical shaded strip indicates the region in which the absolute multiplier effect (squares), divided by relative direct influence of influencers (solid line), yields a relative multiplier effect above unity (dashed line).

initial condition are all regenerated anew, it is possible to associate a distribution of cascade sizes with every choice of parameter settings. The implications for social contagion of different parameter values, and even different models, can then be assessed, either quantitatively or qualitatively, in terms of the properties of the corresponding cascade distributions.

Cascades of any size can and do occur in this model, but an important distinction in what follows is between “local” and “global” cascades. Local cascades affect only a relatively small number of individuals and typically terminate within one or two steps of the initiator. The size of local cascades is therefore determined mostly by the size of an initiator’s immediate circle of influence, not by the size of the network as a whole. Global cascades are the opposite—they affect many individuals, propagate for many steps, and are ultimately constrained only by the size of the population through which they pass. Importantly, global cascades can only occur when the influence network exhibits a “critical mass” of early adopters, which we define here as individuals who adopt after they are exposed to a single adopting neighbor. A critical mass can then be said to exist when sufficiently many early adopters are connected to each other that their subnetwork “percolates” throughout the entire influence network (Watts 2002). Although the critical mass may only occupy a small fraction of the total population, an interesting and subtle consequence of cascade dynamics is that, once it is activated, the remainder of the population subsequently activates as well (Watts 2002), leading to a global cascade. But if the critical mass does not activate, or if it does not exist, then only local cascades are possible. In terms of the diffusion of innovations, the critical mass is therefore what enables a new idea or product to “cross the chasm” from innovation to success (Moore 1999).

### IMPLICATIONS FOR THE INFLUENTIALS HYPOTHESIS

The most obvious way in which an individual can affect the size and likelihood of a cascade is by initiating one. Thus, one way to quantify the relative importance of influencers is to compare the average size of a cascade initiated by an influential to that started by an average member of the population. Figure 3 makes this comparison explicit, both in absolute (3A) and relative (3B) terms. Here we have chosen a value of the average activation threshold  $\phi = 0.18$  for which cascades are possible over some range of the average density  $n_{avg}$  of the influence network, but we have also studied many different choices of  $\phi$ , corresponding to innovations that are, on average, more or less appealing (or, alternatively, populations that are more or less innovative). Although the particular choice of  $\phi$  does affect the probability that a global cascade can take place, it affects influencers and noninfluentials in a similar manner; thus our conclusions regarding the relative effect of influencers are largely independent of  $\phi$ . There are three points to emphasize from this figure, and we discuss each in turn.

First, the size of cascades that can be triggered by single initiators varies enormously depending on the average density  $n_{avg}$  of the influence network. When  $n_{avg}$  is too low, many individuals are highly susceptible to activation, but the network is too poorly connected for influence to propagate far, and thus only a tiny fraction of the network is activated. When  $n_{avg}$  is too high, the opposite applies: the network is highly connected, but individuals now require multiple active neighbors to be activated themselves; thus small initial seeds are unable to grow. Only in the intermediate regime, called the “cascade window” (Watts 2002), can global cascades take place, and in that region both influencers and average individuals are likely to trigger them.

To a first approximation, therefore, the ability of any individual to trigger a cascade depends much more on the global structure of the influence network than on his or her personal degree of influence—that is, if the network permits global cascades, virtually anyone can start one, and if it does not permit global cascades, nobody can. Once again, this statement does not depend on the particular choice of  $\phi$  shown in figure 3, which is purely illustrative. If, for example, we chose some other value of  $\phi$  that also lies within the region in which global cascades occur, the particular curves in figure 3A would shift, but the relationship between them, as shown in figure 3B, would remain qualitatively the same. Furthermore, if we were to choose a value of  $\phi$  such that global cascades were not possible for any other model parameters—corresponding, say, to a highly resilient population or to an unappealing innovation—then the question of who triggers large cascades would obviously be moot: neither influentials nor noninfluentials would be able to do so.

The second point to note, as illustrated in figure 3A, is that when large, multistep cascades do occur, influentials (squares) do tend, on average, to trigger larger cascades (as well as more frequent large cascades) than average individuals (circles). Equivalently, in figure 3B, the relative size of triggered cascades—what Van den Bulte and Joshi (2007) call the “multiplier effect” of influentials (squares)—is always at least one. Thus, although it is clearly not the case that influentials are essential to a global cascade, it is also clear that they have more impact than average individuals. We note, however, that the relative impact of influentials is in most cases modest. For example, when  $n_{\text{avg}} = 3$ , an influential directly influences more than twice as many people ( $n_{\text{inf}} \approx 6$ ) as an average individual, but the multiplier effect (squares) barely exceeds one. Near the lower and upper boundaries of the cascade window, the multiplier effect does clearly exceed one, indicating that influentials are more effective in triggering cascades than ordinary individuals. With the exception of a narrow interval (shaded vertical strip) near the upper boundary, however, the relative multiplier effect (dashed line)—the ratio of the absolute multiplier for influentials to their relative direct influence  $n_{\text{inf}}/n_{\text{avg}}$  (solid line)—is less than one, meaning that the size of the cascades triggered by influentials, although larger than average, is less than proportional to the number of individuals they influence directly.

Taken together, our results provide mixed support for what we have called the influentials hypothesis. The strongest version of the hypothesis—that influentials are in some way essential to diffusion—is clearly not supported, as influentials are neither necessary nor sufficient to trigger large cascades. But the issue of whether or not influentials satisfy some lesser criterion of “importance” beyond their immediate neighborhoods is harder to resolve. Certainly they tend to trigger larger cascades than average, but whether they trigger sufficiently larger cascades to qualify as “important” is a matter of debate. Sometimes, as in the shaded region of figure 3B, the impact of influentials is not only greater

than average but also disproportionately great—a criterion that would seem to serve as a reasonable definition of important. But if that is so, then figure 3 also shows that, under most conditions, influentials are not so important—their global impact is almost always less than proportional to their influence measured locally, and often it is only marginally greater than average.

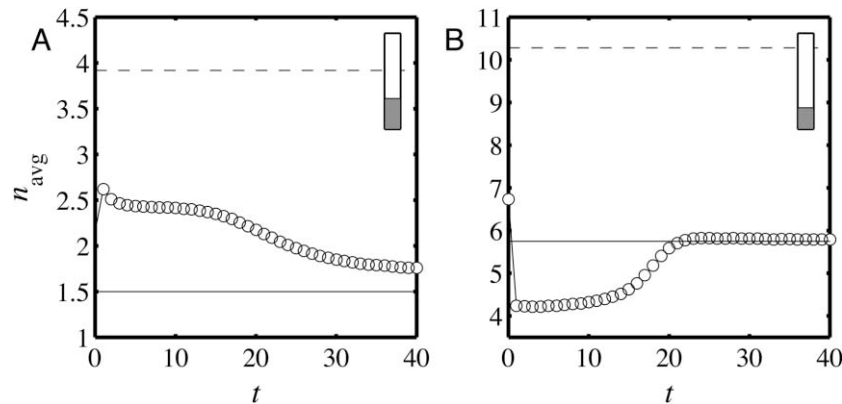
Importance, of course, may be irrelevant. It may be sufficient that influentials are simply more effective than average, even if only marginally so. Given the intense interest that influentials and opinion leaders have attracted ever since they were first hypothesized over 50 years ago, it seems unlikely that marginal importance is all that is being asserted on their behalf. Influentials, by definition, are relatively rare—here they constitute just 10% of the population—thus they are necessarily more difficult to locate than average individuals and possibly more difficult to mobilize also (although that is not necessarily the case). Whether or not the additional impact of influentials justifies paying special attention to them—versus, for example, focusing on some other group, or even recruiting individuals at random—is therefore a matter that will depend, possibly delicately, on the various costs associated with different strategies and the particular details of the influence network.

A second way, however, in which influentials may play a critical role in driving large cascades is as early adopters, who, in our model, make up the critical mass via which local cascades become global. Restricting our attention exclusively to “global” cascades (which in our simulations means cascades that reach more than 1,000 individuals in a population of 10,000), we now consider the extent to which influentials (defined still as the top 10% of the influence distribution) appear as early adopters. Figure 4 shows that, when influence networks are sparse (near the lower boundary of the cascade window), early adopters tend to be more influential than average (4A) but that, when influence networks are dense (i.e., at the upper boundary of the cascade window), the opposite result obtains—early adopters are substantially less influential than average (4B).

That early adopters should be at times more influential than average and at other times less so has been noted previously (Boorman and Levitt 1980; Rogers 1995). However, the origin of the result here is somewhat different from the usual intuition that “central” individuals are highly attuned to group norms and that thus they will tend to adopt early when the group as a whole is progressive and to adopt later when the group is conservative. In our model, no awareness of group norms is necessary or indeed possible. Rather, the principal criterion for an individual to be an early adopter is simply whether or not they can be activated by a single active neighbor. Because the threshold required for this “vulnerability” is inversely proportional to how influential an individual is (Watts 2002), on average, highly influential individuals are less likely to be vulnerable (we will reverse this assumption in the next section). On the other hand, by contrast, the more influential an individual is, the more others they can potentially activate once they themselves are

FIGURE 4

AVERAGE INFLUENCE OF ADOPTERS AS A FUNCTION OF TIME FOR A THRESHOLD MODEL ON RANDOM INFLUENCE NETWORKS



NOTE.—A, plot corresponds to low density,  $n_{\text{avg}} = 1.5$ ; B, plot corresponds to high density,  $n_{\text{avg}} = 5.75$ .

activated. To spread globally, therefore, cascades must strike a compromise, propagating via the most influential of the most easily influenced individuals.

A consequence of this trade-off is that, when  $n_{\text{avg}}$  is low, almost everyone is easy to influence and influential nodes are therefore preferred; but when  $n_{\text{avg}}$  is high, only below-average nodes are sufficiently vulnerable to be activated early on. As figure 4 indicates, however, even when the early adopters are more influential than average, they still do not tend to be influentials. Specifically, we find that, regardless of the density of the influence networks, the average influence of the top 10% of the influence distribution clearly exceeds that of the early adopters, as indicated by the dashed horizontal lines in figures 4A and 4B. At least to the extent that the essential features of influence processes are captured by this model, therefore, large-scale changes in public opinion are not driven by highly influential people who influence everyone else but rather by easily influenced people influencing other easily influenced people.

## VARIATIONS ON THE BASIC INFLUENCE MODEL

Obviously these conclusions, although qualitative in nature, are derived from our analysis of a particular formal model—a model that is, after all, exceedingly simple and that makes a number of important assumptions. One might suspect, therefore, that, if the assumptions of the model were altered in certain respects, our conclusions regarding the importance of influentials would change as well. To address this concern, we now introduce a series of related models that invoke different assumptions with regard to both interpersonal influence and influence networks. It will come as no surprise that these different models can generate at times dramatically different overall dynamics of influence cascades, both quantitatively and qualitatively. We will also

find some additional conditions under which influentials do appear to play important roles, either as initiators or as early adopters. Nevertheless, our conclusion that, under most conditions, cascades are driven by easily influenced, not highly influential, individuals will remain.

## Hyperinfluentials

Perhaps the most obvious objection to the model presented in the previous section is that the influence distribution  $p(n)$  displays relatively little variation around its average, implying that the most influential individuals in our model are, by assumption, not much more influential than average. We would argue that this constraint is, in fact, one of the more reasonable assumptions in our model and that it is quite consistent with the empirical literature on influentials (Burson-Marsteller 2001; Carlson 1965; Coleman et al. 1957; Merton 1968). Recall that opinion leaders are hypothesized to exert interpersonal influence; thus we do not consider media personalities and the like who may have far greater visibility than ordinary individuals but whose influence is transmitted indirectly via media of various forms.

Related to the distinction between personal influence and media influence are the various manifestations of Web-mediated influence, such as that exerted via Web-logs, social networking sites, online forums, and recommender systems. Although individuals can indeed gain considerable exposure for their views by expressing them online—a number of individual bloggers, for example, have gained large followings—the influence of the blogger seems closer to that of a traditional newspaper columnist or professional critic than to that of a trusted confidant or even a casual acquaintance. Thus, although the question of how different forms of influence—including traditional media, Web, and interpersonal influence—compare and interact with each other is indeed an interesting one, it is outside the scope of this article, which

deals only with interpersonal influence. Nevertheless, even in this restricted context, it is interesting to examine how our conclusions would change in a world where, on average, individuals exert roughly the same amount of influence over each other but where the distribution of that influence is more unequal than we have assumed so far.

We have therefore also studied a series of “generalized random networks” (Milo et al. 2004; Newman, Strogatz, and Watts 2001), in which the influence distribution  $p(n)$  can be chosen arbitrarily. Figure 5 shows two examples of influence-degree distributions and their corresponding networks for influence distributions with the same average density  $n_{\text{avg}} = 3$  but with low (5A) and high (5B) variance, respectively. Clearly, the most influential individuals in high-variance networks influence many more of their peers than in low-variance networks of the same average density. For example, in a “low-variance” network of  $N = 10,000$  individuals with  $n_{\text{avg}} = 3$ , the most influential individual influences roughly four times as many others ( $n_{\text{max}} = 12$ ) as an average person, whereas the corresponding individual in a “high-variance” network of the same  $N$  and  $n_{\text{avg}}$  influences roughly 40 times as many ( $n_{\text{max}} = 119$ ) others. Because of their extreme influence, we call these individuals “hyperinfluentials,” where we emphasize that the intensity of influence exerted per relation does not diminish as the

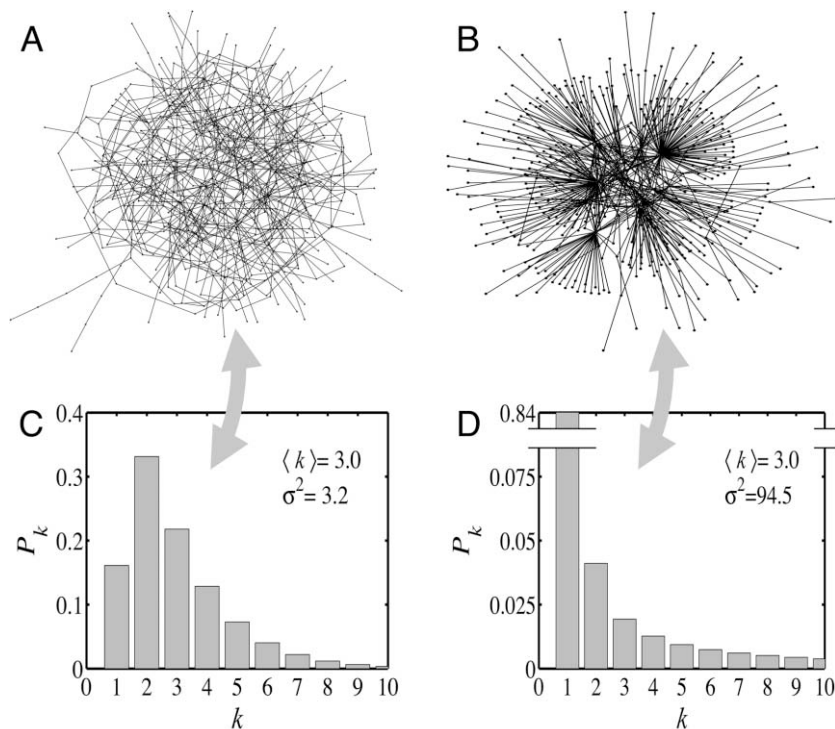
number of influencees increases—a particularly strong assumption in favor of the importance of influentials.

As figure 6 indicates, the presence of hyperinfluentials does indeed affect the size and prevalence of influence cascades. Most strikingly, there are now two regions in figure 6B in which the relative multiplier effect of influentials exceeds one, and these regions are wider than in figure 3B. As one might expect, therefore, the relative impact of hyperinfluentials is greater than that of ordinary influentials, and in some cases, it is impressive—for example, precisely at the lower boundary of the cascade window, influentials trigger cascades that are more than 10 times as large as average (squares, left-hand peak, fig. 6B), even though they directly influence only about six times as many of their peers (solid line). Thus, by assuming that the difference between average and influential individuals is much greater than we did in the previous model, we find that their relative impact is also greater, especially in regions (near the edge of the cascade window) where the conditions for cascades are only marginally satisfied.

Surprisingly, however, the cascades that are triggered by hyperinfluentials are on average less successful than those triggered in low-variance influence networks, even by average individuals (compare squares in fig. 6A with circles in fig. 3A), and the window in which they take place is

FIGURE 5

RANDOM INFLUENCE NETWORKS WITH PRESPECIFIED INFLUENCE DISTRIBUTIONS  $p(n)$

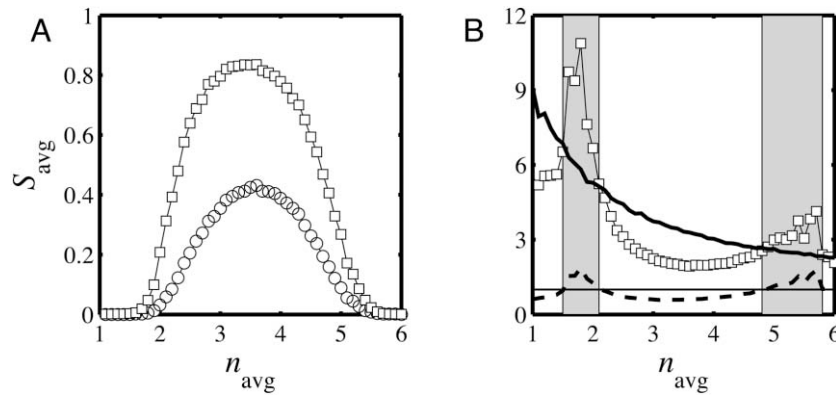


NOTE.—Network on the left exhibits low variance (similar to a standard random network); network on the right exhibits high variance.



FIGURE 6

RELATIVE IMPACT OF INFLUENTIALS AS INITIATORS IN A HIGH-VARIANCE NETWORK



NOTE.—A, Expected cascade size triggered by an influential (squares) and average (circles) individual, respectively, as a function of the density of the influence network. B, The vertical shaded strip indicates the region in which the absolute multiplier effect (squares) for influentials, divided by relative direct influence of influentials (solid line), yields a relative multiplier effect above unity (dashed line).

actually slightly narrower. The reason is that the increased heterogeneity in influence, while advantaging the most influential individuals, disadvantages potential early adopters, who are required for cascades to propagate.

Figure 7 confirms this intuition: early adopters are generally less influential than they are in low-variance networks (fig. 5), and thus they are less able to perpetuate a cascade even though more of them may have been activated initially (i.e., by a hyperinfluential). Thus, the primary requirement that early adopters be vulnerable not only excludes influentials but also impedes the progress of cascades in high-variance networks relative to cascades in their low-variance counterparts. In other words, if one could hypothetically choose between “ordinary” influentials in a low-variance

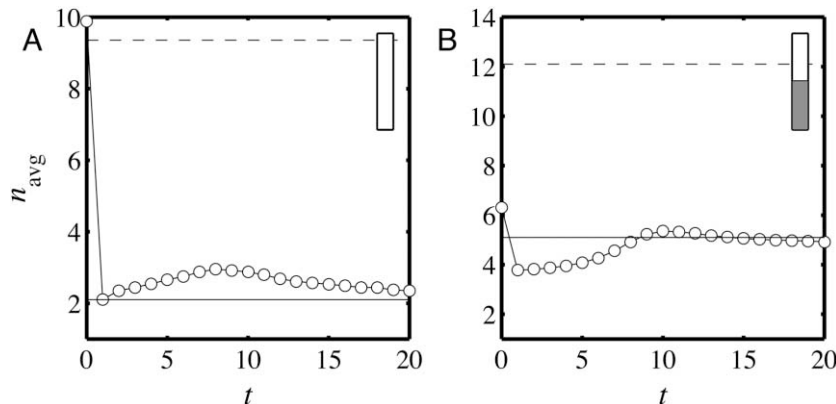
network and hyperinfluentials in a high-variance network, one would, paradoxically, prefer the former.

### Group-Based Networks

Another objection to the basic model is that random networks, whether of low or high variance, are rarely considered good approximations of real social networks. Whereas every acquaintance in a random network is chosen independently of all other acquaintanceships, ties in real social networks exhibit strong interdependencies deriving from various ordering principles such as, for example, social roles (Merton 1957; Nadel 1957), group affiliations (Breiger 1974; Feld 1981), homophily (Lazarsfeld and Merton 1954;

FIGURE 7

AVERAGE INFLUENCE OF ADOPTERS AS A FUNCTION OF TIME IN A HIGH-VARIANCE NETWORK



NOTE.—A, Plot corresponding to low density,  $n_{avg} = 2.1$ . B, plot corresponding to high density,  $n_{avg} = 5.1$ .

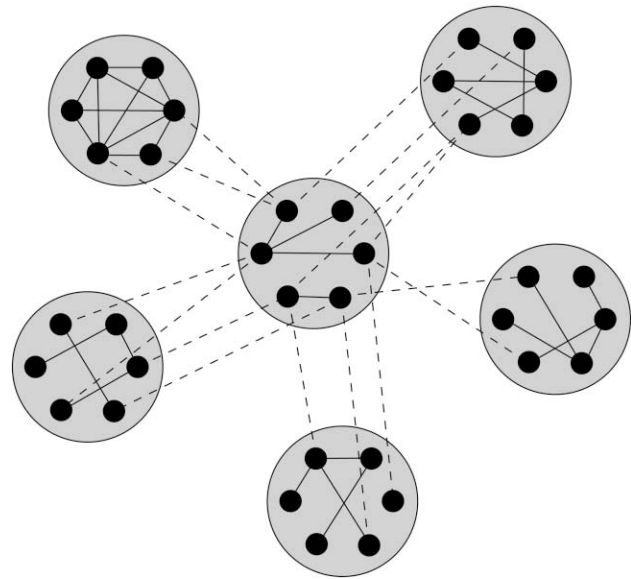
McPherson, Smith-Lovin, and Cook 2001), and triadic closure (Rapoport 1963), all of which have the effect of introducing local structure into the influence network. We have therefore considered a simple form of local structure that generalizes naturally from our baseline random case but which captures the notions (a) that one's acquaintances are more like to influence one another than some individual chosen at random (Watts 1999) and (b) that people have multiple—partly distinct and partly overlapping—groups of acquaintances (Blau and Schwartz 1984; Simmel 1955; Watts, Dodds and Newman 2002).

As illustrated schematically in figure 8, the population of  $N$  individuals is now divided into  $M$  groups of size  $g$  (i.e.,  $N = Mg$ ), and the groups are then randomly paired such that each group, on average, is connected to  $m_{\text{avg}}$  other groups. Each individual  $i$  is then allocated to a single group  $I$ , which is designated his "primary group," and he is then connected to his neighbors as follows: (a) each member of his primary group with probability  $p$  and (b) each member of his  $m_i$  immediately neighboring groups with probability  $q$ . Thus, by varying  $p$  and  $q$ , we can create networks that retain the global connectivity of random networks but that exhibit tunable within- and between-group density. In particular, we consider two kinds of group structure, which we call "integrated" and "concentrated." In integrated networks, each individual has the same probability of knowing a member of one of his neighboring groups as he does of knowing one of his primary group members. In concentrated networks, by contrast, each individual knows at least as many of his primary group members as he does all members of his neighboring groups combined.

As with the introduction of hyperinfluentials, the presence—and even type—of group structure noticeably changes the dynamics of influence cascades. Concentrated networks (bottom row, fig. 9), for example, permit larger cascades than integrated networks (top row), but integrated networks support cascades over a wider interval of  $n_{\text{avg}}$ . Both kinds of group structure, moreover, support cascades over wider intervals than purely random influence networks (compare figs. 9A and 9C with fig. 3A). Nevertheless, the basic conclusions regarding influentials continue to hold—if anything, the addition of group structure appears to diminish their importance. Figures 9B and 9D both indicate that the relative multiplier effect is less than unity everywhere except in a relatively narrow sliver of figure 9B, that is, for integrated networks with low density. In concentrated networks, moreover, for some values of  $n_{\text{avg}}$ , even the absolute multiplier effect (squares) is less than one, meaning that, under some conditions, influentials actually trigger smaller cascades than average individuals. This last result is particularly surprising, but it appears to derive from the observation that influential individuals tend to belong to influential groups; thus, they interact preferentially with other influentials who are, in general, more difficult to influence. Although this tendency, usually known as "assortativity" (Newman 2002) is a consequence of our model's assumptions, it is thought to be characteristic of most real social networks (Newman and Park

FIGURE 8

SCHEMATIC OF THE RANDOM GROUP NETWORK MODEL



2003); thus it is not unreasonable to assume that it arises in influence networks as well.

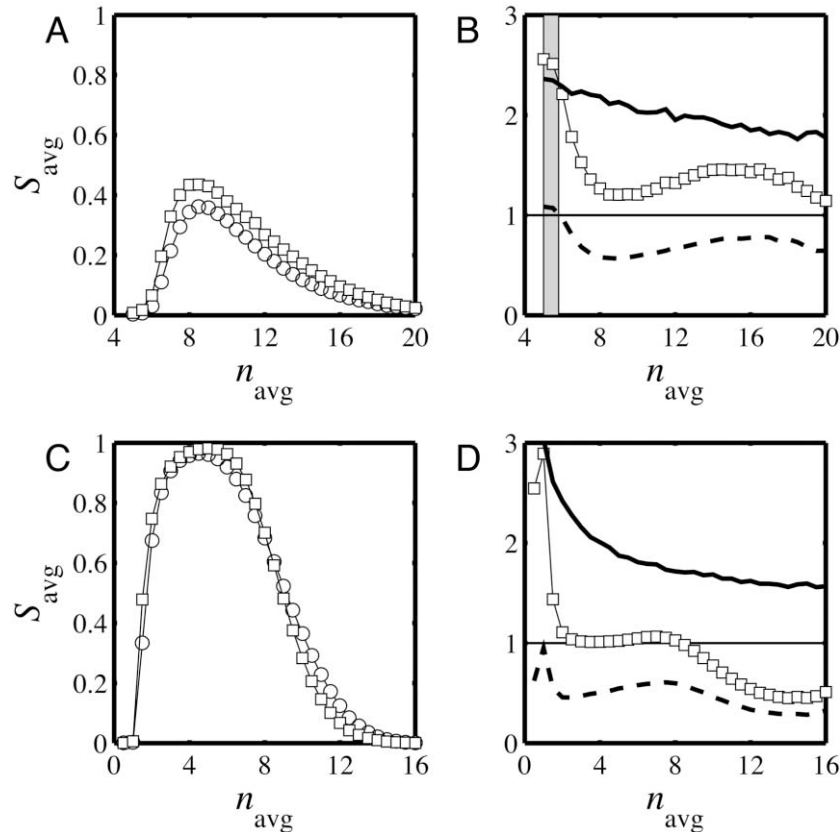
Figure 10, furthermore, shows once again that early adopters are, on average, sometimes above (low density) and sometimes below (high density) the overall average but always below the average influence of influentials (dashed line). The introduction of group structure, in other words, while significantly altering the dynamics of influence cascades, does not much change the ability of influential individuals to trigger or sustain these cascades. If anything group structure appears to further separate the dynamics of interpersonal influence, as observed locally, from the dynamics of global processes like diffusion of innovations or public opinion formation; thus, measurements, say of opinion leadership, that focus on how influential an individual is within their immediate environment are less reliable measures of "importance" than they are in purely random networks.

### Changing the Influence Rule

A final feature of our basic model that one might reasonably question is our particular choice of the influence response function, which assumes that individuals only adopt an innovation when some critical fraction of their neighbors have adopted it. As discussed earlier, threshold rules have been derived for a variety of relevant theoretical scenarios, including coordination and public goods games, so-called network technologies, and adoption decisions in the presence of uncertainty (Lopez-Pintado and Watts 2007). Some limited empirical and experimental evidence also supports the assumption that individuals follow threshold rules when making decisions in the presence of social influence.

FIGURE 9

RELATIVE IMPACT OF INFLUENTIALS AS INITIATORS FOR A THRESHOLD MODEL OF INFLUENCE AND A RANDOM GROUP INFLUENCE NETWORK



NOTE.—Top row reflects “integrated” group structure; bottom row reflects “concentrated” group structure.

In a laboratory experiment in which participants were motivated to anticipate the majority choice of a 24 person group, Latané and L’Herrou (1996) found that individual choice heuristics were well described by a threshold rule with a critical fraction of  $\phi = 1/2$ . Furthermore, Young (2006) has shown that only threshold rules can account for the empirical diffusion curves recorded by Griliches (1957) in his landmark study of the adoption of hybrid corn in the United States in the 1940s.

Nevertheless, threshold rules are only one of many conceivable influence response functions, and in the absence of comprehensive empirical evidence, one must assume that different such functions may arise in different scenarios (Dodds and Watts 2004, 2005). Threshold rules, moreover, carry the additional implication that highly influential individuals are, on average, more difficult to influence—the reason being that an individual who influences many individuals is also influenced by many and thus requires more of his neighbors to adopt than an individual with the same threshold who influences only a few others. Although this implication is certainly plausible, it is not inevitable, and

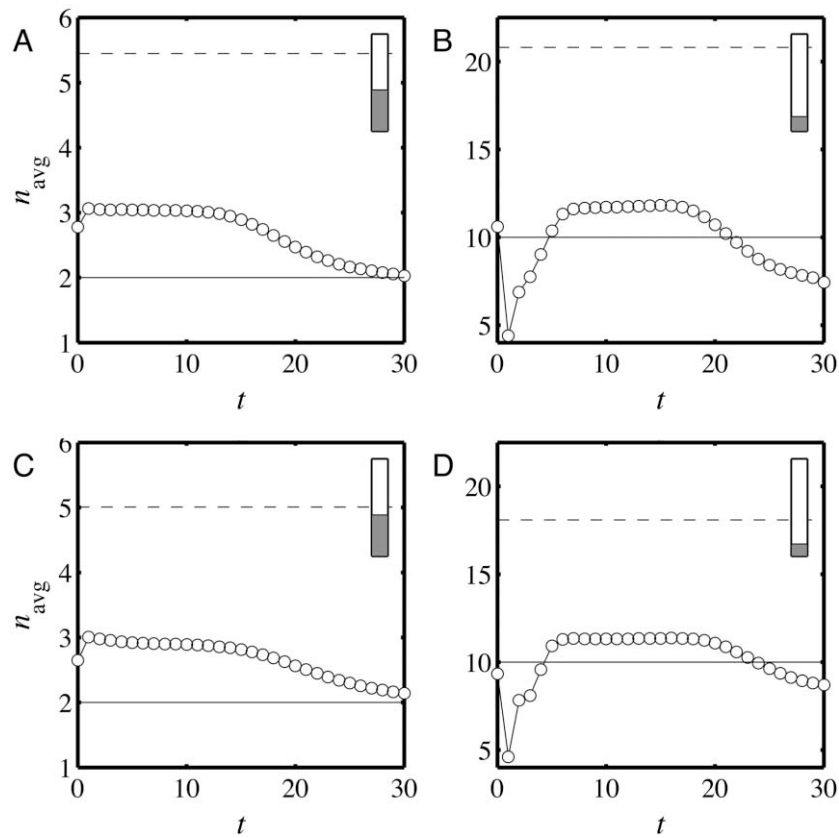
one might wonder how our results regarding the limited impact of influentials would change if highly influential individuals were no more difficult to influence, or perhaps even easier to influence, than average.

To address this concern, we have considered a second canonical type of influence model—known generically as a “SIR” model—several variants of which have appeared in the diffusion of innovations literature (Coleman et al. 1957; Dodson and Muller 1978; Lehmann and Esteban-Bravo 2006; Van den Bulte and Joshi 2007), including the Bass (1969) model, and have been studied extensively in mathematical epidemiology (Anderson and May 1991). In the SIR model, individuals can occupy one of three states—“susceptible” (S), “infected” (I), and “recovered” (R)—where susceptible individuals become infected with probability  $\beta$  (the infectiousness) when they encounter an infective and where they subsequently recover at rate  $r$ .

Although the SIR model, like the threshold model, treats interpersonal influence as a form of social contagion, the two models incorporate fundamentally different assumptions about the contagious process itself (Dodds and Watts

FIGURE 10

AVERAGE INFLUENCE OF ADOPTERS AS A FUNCTION OF TIME FOR RANDOM-GROUP INFLUENCE NETWORKS



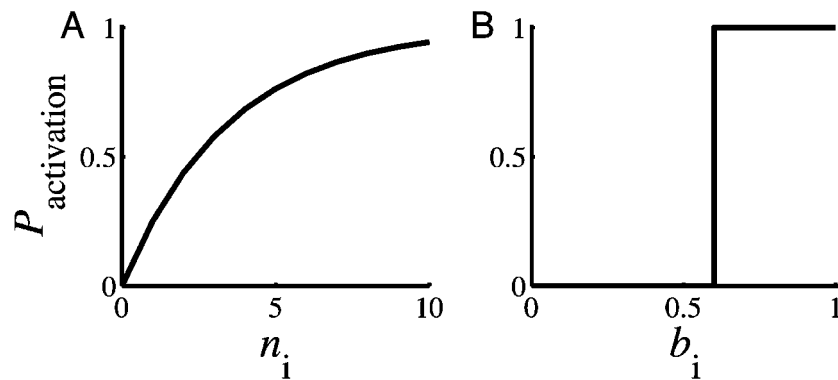
NOTE.—Left-hand column plots are for low-density networks,  $n_{\text{avg}} = 2$ ; right-hand column plots are for high-density ones,  $n_{\text{avg}} = 10$ ; top row corresponds to integrated groups; bottom row corresponds to concentrated groups.

2005). The threshold model assumes, in effect, a cost-benefit calculation on the part of the adopter, and therefore requires that decision makers can, in effect, remember a sequence of observations over time. SIR models, by contrast, are “pure” contagion models in the sense that every contact between an infective and a susceptible is treated independently of any other. Mathematically, the models are very different as well. As illustrated in figure 11, the influence response function for the SIR model is a concave function that increases with the absolute number of “infected” contacts, whereas the threshold response function is at first convex and then concave and depends on the infected fraction. The difference is central to our argument because, in the SIR model, the more influential an individual is, the more easily influenced he or she is as well, which is the opposite of the threshold case. It has recently been shown, in fact, that SIR and threshold models are extreme cases on a continuum of possible contagion models (Dodds and Watts 2004, 2005); thus, by testing the SIR model, we effectively span a wide range of possible assumptions about binary decision models with positive externalities.

Figure 12 shows, as might be expected, that the different choice of influence model has a dramatic impact on the dynamics of social influence propagation, both for low-variance (top row) and high-variance (bottom row) networks, where the latter contains what we have called hyperinfluentials. Most strikingly, there is no longer an upper bound to the cascade window. Because highly influential individuals are now more, not less, susceptible to influence themselves, as the networks become more connected, cascades become larger and also more frequent. It is perhaps surprising, therefore, that under such dramatically different circumstances, our conclusions with respect to influentials are largely unchanged. As before, influentials tend to trigger marginally larger cascades than average individuals, but their relative impact remains limited—in fact, even more limited for the SIR model than it is for the threshold model. In particular, unlike with other models, here the relative multiplier effect is always less than one, even, surprisingly, when hyperinfluentials are present (fig. 12D). The reason, however, is not that influentials trigger smaller cascades than they do in the

FIGURE 11

INFLUENCE RESPONSE FUNCTIONS FOR (A) THE SIR MODEL AND (B) THE DETERMINISTIC THRESHOLD RULE



NOTE.—Each function reflects the probability of choosing alternative B as a function of the number  $n_i$  or as a fraction  $b_i$  of others choosing B, respectively.

threshold model but rather that average individuals trigger larger ones.

Another qualitative difference between the SIR and threshold models is that early adopters are consistently more influential than average in both low- and high-density regimes, as shown in figure 13 (again, for low- and high-variance networks in the top row and the bottom row, respectively). The reason, once again, is that more influential individuals are more, not less, susceptible to influence themselves; thus, there is no trade-off between influence and influenceability. Given this difference, one might expect that, in the SIR model, the population of early adopters will include more influentials. As figure 13 shows, this intuition is not supported for the case of “ordinary” influentials (top row) but is supported for hyperinfluentials when the corresponding network is also sufficiently sparse: the first two “generations” of early adopters in figure 13C are clearly influentials. Thus, although hyperinfluentials do not appear to be disproportionately effective in triggering large cascades in the SIR model, their combination of greater influence and greater influenceability renders them important as very early adopters, at least under some conditions.

In summary, the series of models we have considered offer some reasonably clear guidelines with regard to what conditions must be satisfied in order for influentials to play an important role in processes like public opinion formation. First, “ordinary” influentials of the kind considered in low-variance networks appear to be important as initiators of large cascades where the threshold rule is in effect and in conditions where these cascades are only marginally possible (fig. 3B). They do not, however, play important roles as initiators under most conditions of the threshold model, under any conditions as early adopters, or when the SIR model is in effect. Second, “hyperinfluentials” of the kind that arise in high-variance networks are important as initiators under a wider range of conditions than ordinary influentials but still only when the threshold rule is in effect. Third, when the SIR rule is in effect, hyperinfluentials play

an important role as early adopters, when networks are sufficiently sparse, but not as initiators. Finally, group structure appears to generally impede the effectiveness of influentials both as initiators and early adopters.

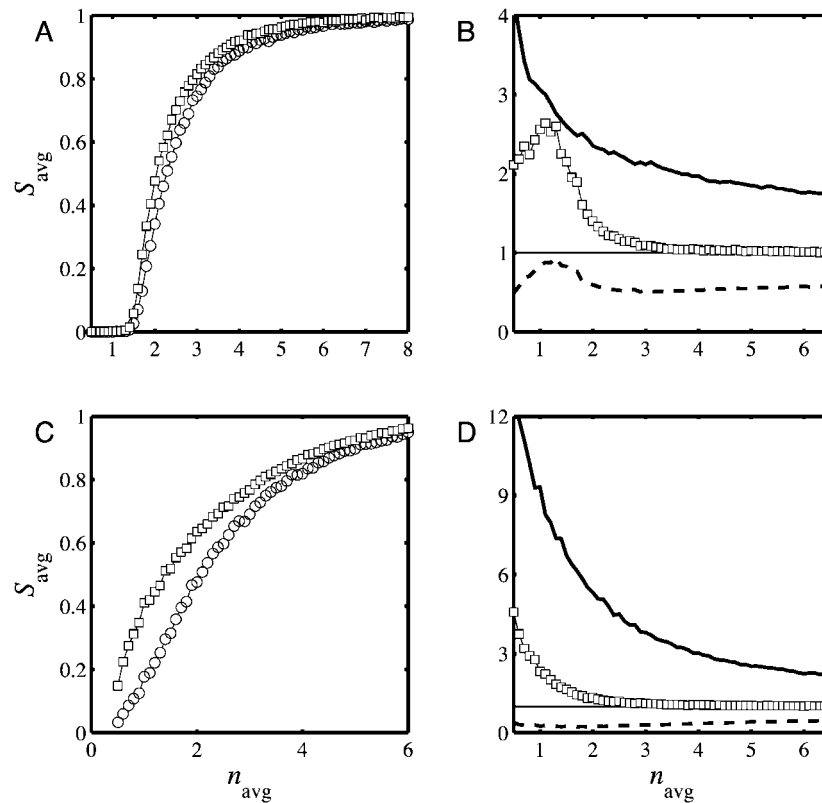
## DISCUSSION

Whether these results should be regarded as undermining what we have called the influentials hypothesis or as supporting it is ultimately an empirical question. Our main point, in fact, is not so much that the influentials hypothesis is either right or wrong but that its microfoundations, by which we mean the details of who influences whom and how, require very careful articulation in order for its validity to be meaningfully assessed. Whether stated explicitly or not, any claim to the effect that influentials are important necessarily makes a number of assumptions regarding the nature of interpersonal influence, the structure of influence networks, and even what is meant by “important.” Regardless of which particular parameter values and even influence models future empirical work turns out to support, therefore, the primary contribution of simulation studies like the one we have conducted is to highlight the importance of specifying these assumptions precisely and to focus empirical attention on those that seem most relevant.

Looking over the full range of models and assumptions that we have considered, however, we would argue that our study also makes a second, albeit more tentative, contribution, which is that, under most of these conditions, influentials are less important than is generally supposed, either as initiators of large cascades or as early adopters. This statement, we emphasize, does not mean that influential individuals do not exist or that they never display the kind of importance that is implied by the influentials hypothesis. Nevertheless, the conditions we have identified under which influential individuals can play important roles are neither particularly common nor do they seem especially likely to arise in many real situations. The existence of hyperinfluen-

FIGURE 12

RELATIVE IMPACT OF INFLUENTIALS AS INITIATORS FOR A SIR MODEL OF INFLUENCE AND A RANDOM INFLUENCE NETWORK



NOTE.—Compare to fig. 3. Symbols and line styles are as in fig. 3. Top row corresponds to low-variance networks, and bottom row corresponds to high-variance networks.

tials, for example, seems to us more a theoretical possibility than an empirical reality—we are not aware of any empirical studies in which individuals have been shown to influence over 100 others directly—especially if one also requires them to be easier to influence than average individuals.

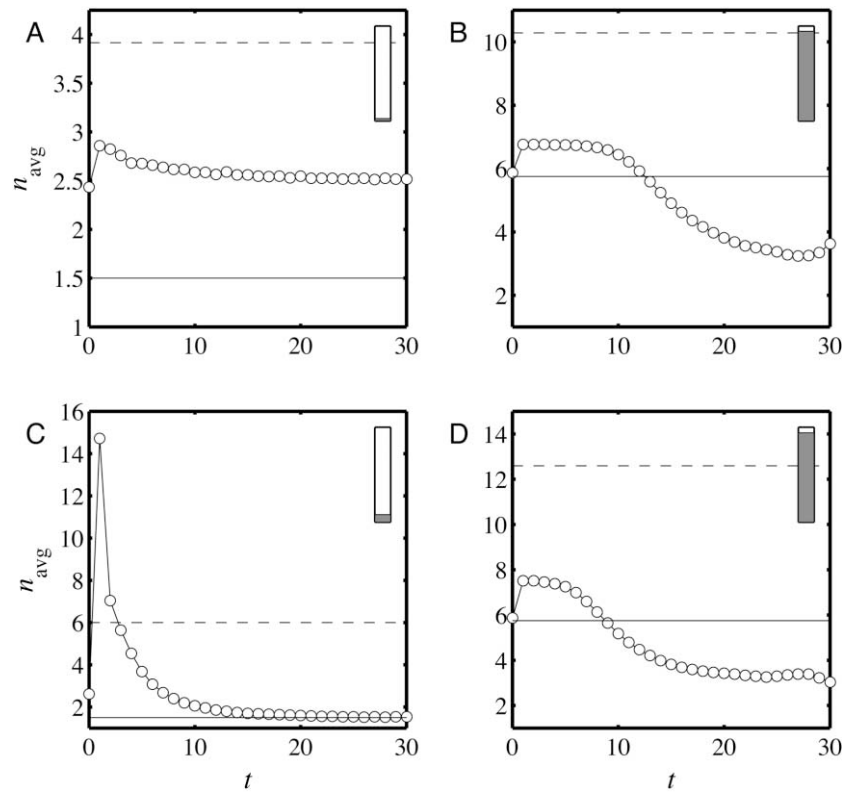
In fact, while any assertion regarding the lack of importance of influentials is necessarily speculative, based on our results, we would go as far as to suggest that in focusing on the properties of a few “special” individuals, the influentials hypothesis is in some important respects a misleading model for social change. Under most conditions, we would argue, cascades do not succeed because of a few highly influential individuals influencing everyone else but rather on account of a critical mass of easily influenced individuals influencing other easy-to-influence people. In our models, influentials have a greater than average chance of triggering this critical mass, when it exists, but only modestly greater, and usually not even proportional to the number of people they influence directly. They may also participate in the critical mass, especially when they are simultaneously hyperinfluential and easily influenced, but under most conditions they do not. Thus, to the extent that particular indi-

viduals appear, after the fact, to have been disproportionately responsible for initiating a large cascade or sustaining it in its early stages, the identities and even characteristics of those individuals are liable to be accidents of timing and location, not evidence of any special capabilities or superior influence.

Although this claim may seem implausible in light of received wisdom, it has numerous analogues in natural systems. Some forest fires, for example, are many times larger than average; yet no one would claim that the size of a forest fire can be in any way attributed to the exceptional properties of the spark that ignited it or the size of the tree that was the first to burn. Major forest fires require a conspiracy of wind, temperature, low humidity, and combustible fuel that extends over large tracts of land. Just as for large cascades in social influence networks, when the right global combination of conditions exists, any spark will do; when it does not, none will suffice. Avalanches, earthquakes, and other natural disasters all share a similar ambiguity with respect to initial conditions—in all these cases, very similar initial disturbances can, under some conditions, lead to dra-

FIGURE 13

AVERAGE INFLUENCE OF ADOPTERS AS A FUNCTION OF TIME FOR A SIR MODEL



NOTE.—Plot A corresponds to low-density ( $n_{avg} = 1.5$ ) and plot B corresponds to high-density ( $n_{avg} = 5.75$ ) random influence networks.

matically different outcomes; under other conditions, very different disturbances can yield indistinguishable results.

Analogous results in human systems can also be inferred from the literature on so-called information cascades (Bikhchandani et al. 1992, 1998; De Vany and Walls 1996) in which small, possibly random, fluctuations at critical junctures cause even large groups to become locked into a particular collective choice. Because the outcome of the cascade depends to such a great extent on the choices of a small number of individuals who arrive at the right time, anyone observing the final outcome might be tempted to label them as influentials. Yet models of information cascades, as well as human subjects experiments that have been designed to test the models (Anderson and Holt 1997; Kubler and Weizsacker 2004), are explicitly constructed such that there is nothing special about those individuals, either in terms of their personal characteristics or in their ability to influence others. Thus, whatever influence these individuals exert on the collective outcome is an accidental consequence of their randomly assigned position in the queue—a result that has been demonstrated recently for more complicated and realistic scenarios as well (Salganik et al. 2006).

Nevertheless, anytime some notable social change is recognized, whether it be a grassroots cultural fad, a successful

marketing campaign, or a dramatic drop in crime rates, it is tempting to trace the phenomenon to the individuals who “started it” and to conclude that their actions or behavior “caused” the events that subsequently took place. Indeed, because the outcome is already known, it is always possible to construct what looks like a causal story by picking out some of the defining details of the individuals in question—even when success is completely random (Taleb 2001). Likewise, it is tempting to assert that these individuals must have been special in some way—otherwise, how could the striking event that we now know happened have come to pass? Just because the outcome is striking, however, does not on its own imply that there is anything correspondingly special about the characteristics of the individuals involved (Goldenberg, Lehmann, and Mazursky 2001) or that their participation was either a necessary or sufficient condition for a change of the kind that occurred to have taken place (Lieberson 1991). Although simulation models of the kind we have studied here cannot ultimately resolve such debates, they can contribute to them by prompting us to question our intuitive understanding of social processes, by suggesting alternative explanations, and by pointing out possible directions for empirical studies.

## REFERENCES

- Anderson, Lisa R. and Charles A. Holt (1997), "Information Cascades in the Laboratory," *American Economic Review*, 87, 847-62.
- Anderson, Roy M. and Robert M. May (1991), *Infectious Diseases of Humans*, Oxford: Oxford University Press.
- Arndt, John (1967), "Role of Product-Related Conversations in the Diffusion of a New Product," *Journal of Marketing Research*, 4, 291-95.
- Arthur, W. Brian (1994), "Inductive Reasoning and Bounded Rationality," *American Economic Review*, 84, 406-11.
- Bailey, N. T. J. (1975), *The Mathematical Theory of Infectious Diseases and Its Applications*, New York: Hafner.
- Banerjee, Abhijit V. (1992), "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107, 797-817.
- Barabasi, Albert-Laszlo (2002), *Linked: The New Science of Networks*, Cambridge, MA: Perseus.
- Bass, Frank M. (1969), "A New Product Growth for Model Consumer Durables," *Management Science*, 15, 215-27.
- Bernheim, B. Douglas (1994), "A Theory of Conformity," *Journal of Political Economy*, 102, 841-77.
- Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch (1992), "A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades," *Journal of Political Economy*, 100, 992-1026.
- — — (1998), "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades," *Journal of Economic Perspectives*, 12, 151-70.
- Bineham, J. L. (1988), "A Historical Account of the Hypodermic Model in Mass Communication," *Communication Monographs*, 55, 230-46.
- Blau, Peter M. and J. E. Schwartz (1984), *Crosscutting Social Circles*, Orlando, FL: Academic Press.
- Blume, Lawrence and Steven N. Durlauf (2003), "Equilibrium Concepts for Social Interaction Models," *International Game Theory Review*, 5, 193-209.
- Bond, Rod and Peter B. Smith (1996), "Culture and Conformity: A Meta-analysis of Studies Using Asch's (1952b, 1956) Line Judgment Task," *Psychological Bulletin*, 119, 111-37.
- Boorman, Scott A. and Paul R. Levitt (1980), *The Genetics of Altruism*, New York: Academic Press.
- Borghesi, Christian and Jean-Philippe Bouchaud (forthcoming), "Of Songs and Men: A Model for Multiple Choice with Herding," *Quality and Quantity*, <http://www.xxx.lanl.gov/pdf/physics/0606224>.
- Breiger, R. (1974), "The Duality of Persons and Groups," *Social Forces*, 53, 181-90.
- Brock, William A. and Steven N. Durlauf (2001), "Discrete Choice with Social Interactions," *Review of Economic Studies*, 68, 235-60.
- Brown, Jacqueline Johnson and Peter H. Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research*, 14, 350-62.
- Burson-Marsteller (2001), "The Power of Online Influencers," Report, Burson-Marsteller, New York, <http://www.efluentials.com/documents/Brochure.pdf>.
- Burt, Ronald S. (1987), "Social Contagion and Innovation: Cohesion versus Structural Equivalence," *American Journal of Sociology*, 92, 1287-335.
- — — (1999), "The Social Capital of Opinion Leaders," *Annals of the American Academy of Political and Social Science*, 566, 37-54.
- Carlson, Richard O. (1965), *Adoption of Educational Innovations*, Eugene: University of Oregon.
- Chan, K. K. and S. Misra. (1990), "Characteristics of the Opinion Leader: A New Dimension," *Journal of Advertising*, 19, 53-60.
- Cialdini, Robert B. (2001), *Influence: Science and Practice*, Needham Heights, MA: Allyn & Bacon.
- Cialdini, Robert B. and J. Noah Goldstein (2004), "Social Influence: Compliance and Conformity," *Annual Review of Psychology*, 55, 591-621.
- Coleman, James Samuel, Elihu Katz, and Herbert Menzel (1957), "The Diffusion of an Innovation among Physicians," *Sociometry*, 20, 253-70.
- — — (1966), *Medical Innovation: A Diffusion Study*, Indianapolis: Bobbs-Merrill.
- Coulter, Robin A., Lawrence F. Feick, and Linda L. Price (2002), "Changing Faces: Cosmetics Opinion Leadership among Women in the New Hungary," *European Journal of Marketing*, 36, 1287-1308.
- De Vany, Arthur and W. David Walls (1996), "Bose-Einstein Dynamics and Adaptive Contracting in the Motion Picture Industry," *Economic Journal*, 106, 1493-1514.
- Dodds, Peter Sheridan and Duncan J. Watts (2004), "Universal Behavior in a Generalized Model of Contagion," *Physical Review Letters*, 92 (21), no. 218701.
- — — (2005), "A Generalized Model of Social and Biological Contagion," *Journal of Theoretical Biology*, 232, 587-604.
- Dodson, Joe A. and Eitan Muller (1978), "Models of New Product Diffusion through Advertising and Word-of-Mouth," *Management Science*, 24, 1568-78.
- Doumit, G., M. Gattellari, J. Grimshaw, and M. A. O'Brien (2007), "Local Opinion Leaders: Effects on Professional Practice and Health Care Outcomes," in Cochrane Database of Systematic Reviews, Cochrane Library, <http://www.mrw.interscience.wiley.com/cochrane/clsysrev/articles/CD000125/frame.html>.
- Durlauf, Steven N. (2001), "A Framework for the Study of Individual Behavior and Social Interactions," *Sociological Methodology*, 31, 47-87.
- Feld, Scott L. (1981), "The Focused Organization of Social Ties," *American Journal of Sociology*, 86, 1015-35.
- Gitlin, Todd (1978), "Media Sociology: The Dominant Paradigm," *Theory and Society*, 6, 205-53.
- Gladwell, Malcolm (2000), *The Tipping Point: How Little Things Can Make a Big Difference*, New York: Little, Brown.
- Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*, 23, 545-60.
- Goldenberg, Jacob, Donald R. Lehmann, and David Mazursky (2001), "The Idea Itself and the Circumstances of Its Emergence as Predictors of New Product Success," *Management Science*, 47, 69-84.
- Goldenberg, Jacob, Barak Libai, and Eitan Muller (2001), "Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth," *Marketing Letters*, 12 (3), 211-24.
- Granovetter, M. S. (1978), "Threshold Models of Collective Behavior," *American Journal of Sociology*, 83, 1420-43.
- Grewal, R., R. Mehta, and F. R. Kardes (2000), "The Role of the Social-Identity Function of Attitudes in Consumer Innovativeness and Opinion Leadership," *Journal of Economic Psychology*, 21, 233-52.
- Griliches, Zvi (1957), "Hybrid Corn: An Exploration of the Economics of Technological Change," *Econometrica*, 25, 501-22.
- Hedstrom, Peter (1998), "Rational Imitation," in *Social Mecha-*



- nisms: An Analytical Approach to Social Theory*, ed. Peter Hedstrom and Richard Swedberg, Cambridge: Cambridge University Press, 306–27.
- Katz, Elihu and Paul Felix Lazarsfeld (1955), *Personal Influence; the Part Played by People in the Flow of Mass Communications*, Glencoe, IL: Free Press.
- Kearns, Michael, Siddharth Suri, and Nick Montfort (2006), “An Experimental Study of the Coloring Problem on Human Subject Networks,” *Science*, 313, 824–27.
- Keller, Ed and Jon Berry (2003), *The Influentials: One American in Ten Tells the Other Nine How to Vote, Where to Eat, and What to Buy*, New York: Free Press.
- Kelly, J. A., J. S. St. Lawrence, Y. E. Diaz, L. Y. Stevenson, A. C. Hauth, T. L. Brasfield, S. C. Kalichman, J. E. Smith, and M. E. Andrew (1991), “HIV Risk Behavior Reduction Following Intervention with Key Opinion Leaders of Population: An Experimental Analysis,” *American Journal of Public Health*, 1, 168–71.
- Kubler, Dorothea and Georg Weizsacker (2004), “Limited Depth of Reasoning and Failure of Cascade Formation in the Laboratory,” *Review of Economic Studies*, 71, 425–41.
- Kuran, Timur (1991), “Now out of Never: The Element of Surprise in the East European Revolution of 1989,” *World Politics*, 44, 7–48.
- Latané, Bibb and Todd L’Herrou (1996), “Spatial Clustering in the Conformity Game: Dynamic Social Impact in Electronic Groups,” *Journal of Personality and Social Psychology*, 70, 1218–30.
- Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet (1968), *The People’s Choice: How the Voter Makes Up His Mind in a Presidential Campaign*, New York: Columbia University Press.
- Lazarsfeld, Paul, and Robert Merton (1954), “Friendship as Social Process: A Substantive and Methodological Analysis,” in *Freedom and Control in Modern Society*, ed. Monroe Berger, Theodore Abel, and Charles Page, New York: Van Nostrand, 18–66.
- Lehmann, Donald and Mercedes Esteban-Bravo (2006), “When Giving Some Away Makes Sense to Jump-Start the Diffusion Process,” *Marketing Letters*, 17 (4), 243–54.
- Leibenstein, H. (1950), “Bandwagon, Snob, and Veblen Effects in the Theory of Consumers’ Demand,” *Quarterly Journal of Economics*, 64, 183–207.
- Leskovec, Jurij, Lada A. Adamic, and Bernardo A. Huberman (2007), “The Dynamics of Viral Marketing,” *ACM Transactions on the Web*, 1 (1), forthcoming.
- Lieberman, Stanley (1991), “Small N’s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases,” *Social Forces*, 70, 307–20.
- Liebowitz, Stanley J. and Stephen E. Margolis (1998), “Network Effects and Externalities,” in *The New Palgrave’s Dictionary of Economics and the Law*, ed. Peter Newman, New York: Macmillan, 671–75.
- Lin, Nan (1973), *The Study of Human Communication*, Indianapolis: Bobbs-Merrill.
- Lopez-Pintado, Dunia and Duncan J. Watts (2007), “Social Influence, Binary Decisions, and Collective Dynamics,” Working paper, Institute for Social and Economic Research and Policy, Columbia University.
- Mayzlin, Dina (2002), “The Influence of Social Networks on the Effectiveness of Promotional Strategies,” working paper, Yale School of Management.
- McPherson, Miller, Lynn Smith-Lovin, and James M. Cook (2001), “Birds of a Feather: Homophily in Social Networks,” *Annual Review of Sociology*, 27, 415–44.
- Menzel, Herbert and Elihu Katz (1955), “Social Relations and Innovation in the Medical Profession: The Epidemiology of a New Drug,” *Public Opinion Quarterly*, 19, 337–52.
- Merton, Robert K. (1957), “The Role-Set: Problems in Sociological Theory,” *British Journal of Sociology*, 8, 106–20.
- — — (1968), “Patterns of Influence: Local and Cosmopolitan Influentials,” in *Social Theory and Social Structure*, ed. Robert K. Merton, New York: Free Press, 441–74.
- Milo, R., N. Kashtan, S. Itzkovitz, M. E. J. Newman, and U. Alon (2004), “On the Uniform Generation of Random Graphs with Prescribed Degree Sequences,” <http://arxiv.org/abs/cond-mat/0312028>.
- Moore, Geoffrey A. (1999), *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*, New York: Harper Business.
- Moore, K. A., R. H. Peters, H. A. Hills, J. B. LeVasseur, A. R. Rich, W. M. Hunt, M. S. Young, and T. W. Valente (2004), “Characteristics of Opinion Leaders in Substance Abuse Treatment Agencies,” *American Journal of Drug and Alcohol Abuse*, 30, 187–203.
- Morris, S. (2000), “Contagion,” *Review of Economic Studies*, 67, 57–78.
- Myers, James H. and Thomas S. Robertson (1972), “Dimensions of Opinion Leadership,” *Journal of Marketing Research*, 9, 41–46.
- Nadel, F. S. (1957), *Theory of Social Structure*, Glencoe, IL: Free Press.
- Newman, M. E. J. (2002), “Assortative Mixing in Networks,” *Physical Review Letters*, 89, no. 208701.
- — — (2003), “The Structure and Function of Complex Networks,” *SIAM Review*, 45, 167–256.
- Newman, M. E. J. and J. Park (2003), “Why Social Networks Are Different from Other Types of Networks,” *Physical Review E*, 68, no. 036122.
- Newman, M. E. J., S. H. Strogatz, and D. J. Watts, (2001), “Random Graphs with Arbitrary Degree Distributions and Their Applications,” *Physical Review E*, 6402, no. 026118.
- Nisbet, Erik C. (2006), “The Engagement Model of Opinion Leadership: Testing Validity within a European Context,” *International Journal of Public Opinion Research*, 18, 3–30.
- Oliver, Pamela E. and Gerald Marwell (1985), “A Theory of the Critical Mass. I. Interdependence, Group Heterogeneity, and the Production of Collective Action,” *American Journal of Sociology*, 91, 522–56.
- Rand, Paul M. (2004), “Identifying and Reaching Influencers,” <http://www.marketingpower.com/content20476.php>.
- Rapoport, A. (1963), “Mathematical Models of Social Interaction,” in *Handbook of Mathematical Psychology*, ed. R. D. Luce, R. R. Bush, and E. Galanter. New York: Wiley, 493–579.
- Robinson, John P. (1976), “Interpersonal Influence in Election Campaigns: Two Step-Flow Hypotheses,” *Public Opinion Quarterly*, 40, 304–19.
- Roch, C. H. (2005), “The Dual Roots of Opinion Leadership,” *Journal of Politics*, 67, 110–31.
- Rogers, Everett M. (1995), *Diffusion of Innovations*, New York: Free Press.
- Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts, (2006), “Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market,” *Science*, 311, 854–56.
- Schelling, Thomas C. (1973), “Hockey Helmets, Concealed Weapons,

- ons, and Daylight Saving: A Study of Binary Choices with Externalities," *Journal of Conflict Resolution*, 17, 381–428.
- — — (1978), *Micromotives and Macrobehavior*, New York: Norton.
- Shrager, J., T. Hogg, and B. A. Huberman (1987), "Observation of Phase-Transitions in Spreading Activation Networks," *Science*, 236, 1092–94.
- Simmel, Georg (1955), *Conflict and the Web of Group Affiliations*, Glencoe, IL: Free Press.
- Solomonoff, R. and A. Rapoport (1951), "Connectivity of Random Nets," *Bulletin of Mathematical Biophysics*, 13, 107–17.
- Soumerai, Stephen B., Thomas J. McLaughlin, Jerry H. Gurwitz, Edward Guadagnoli, Paul J. Hauptman, Catherine Borbas, Nora Morris, Barbara McLaughlin, Xiaoming Gao, Donald J. Willison, Richard Asinger, and Fredarick Gobel (1998), "Effect of Local Medical Opinion Leaders on Quality of Care for Acute Myocardial Infarction: A Randomized Controlled Trial," *JAMA* 279 (17): 1358–63.
- Taleb, Nassim Nicholas (2001), *Foiled by Randomness*, New York: Norton.
- Troldahl, Verling C. and Robert Van Dam (1965), "Face-to-Face Communication about Major Topics in the News," *Public Opinion Quarterly*, 29, 626–34.
- Valente, Thomas W. (1995), *Network Models of the Diffusion of Innovations*, Cresskill, NJ: Hampton.
- Van den Bulte, Christophe and Yogesh V. Joshi (2007), "New Product Diffusion with Influentials and Imitators," *Marketing Science*, forthcoming.
- Vernette, E. (2004), "Targeting Women's Clothing Fashion Opinion Leaders in Media Planning: An Application for Magazines," *Journal of Advertising Research*, 44, 90–107.
- Wasserman, Stanley and Katherine Faust (1994), *Social Network Analysis: Methods and Applications*, Cambridge: Cambridge University Press.
- Watts, Duncan J. (1999), "Networks, Dynamics, and the Small World Phenomenon," *American Journal of Sociology* 105, 493–527.
- — — (2002), "A Simple Model of Information Cascades on Random Networks," *Proceedings of the National Academy of Science* 99, 5766–71.
- — — (2003), *Six Degrees: The Science of a Connected Age*. New York: Norton.
- — — (2004), "The 'New' Science of Networks," *Annual Review of Sociology*, 30, 243–70.
- Watts, Duncan J., Peter Sheridan Dodds, and Mark E. J. Newman (2002), "Identity and Search in Social Networks," *Science*, 296, 1302–5.
- Weimann, Gabriel (1991), "The Influentials: Back to the Concept of Opinion Leaders?" *Public Opinion Quarterly*, 55, 267–79.
- — — (1994), *The Influentials: People Who Influence People*, Albany: SUNY Press.
- Young, H. Peyton (1996), "The Economics of Convention," *Journal of Economic Perspectives*, 10, 105–22.
- — — (2006), "The Spread of Innovations through Social Learning," Working paper, Center on Social and Economic Dynamics, Brookings Institution.