www.sciencemag.org/cgi/content/full/329/5996/1194/DC1

Supporting Online Material for

The Spread of Behavior in an Online Social Network Experiment

Damon Centola

E-mail: dcentola@mit.edu

Published 3 September 2010, *Science* **329**, 1194 (2010)
DOI: 10.1126/science.1185231

This PDF file includes:

Supporting Online Materials for

The Spread of Behavior in an Online Social Network Experiment

This file includes:

Experimental Design

As described in the text, each trial of the study consisted of a pair of networks, one of each condition – a *clustered lattice network* condition and a *random network* condition – between which subjects were randomly assigned. The schema for this design is shown in Figure S1. Once subjects were assigned to a network condition, they were randomly assigned to one node in the network. In each condition, a single, randomly selected node was not filled with a subject – this node was used as the "seed node" for initiating the diffusion dynamics once the network was populated. In each trial, both networks had the same degree (each node had the same number of neighbors as every other node), and the same overall size (each network had the same total number of nodes). Within each trial, the only difference between conditions was the topological structure of the ties that connected the participants. Consequently, any difference in the dynamics of diffusion between conditions was due to the effects of network topology on the spread of behavior. Six independent trials of the study were run. Across all six trials, 50% of subjects were enrolled in the clustered lattice network condition and 50% were enrolled in the random network condition.

Diffusion dynamics were initiated by sending an email signal from the randomly chosen seed node to its neighbors in the social network, inviting them to register for a health

forum website that offered access and rating tools for online health resources. If any of these subjects registered for the site, their network neighbors (i.e., "health buddies"), in turn, received an email message inviting them to adopt the behavior. If a subject had multiple network neighbors who adopted the behavior, then she would receive multiple email signals, one from each neighbor. The sequence, timing, and number of signals that a subject received was entirely determined by the adoption pattern of the members of the social network. All email signals were sent from the "Healthy Lifestyle Network," and subjects had no access to identifying information about who their health buddies were, or how to contact them directly.

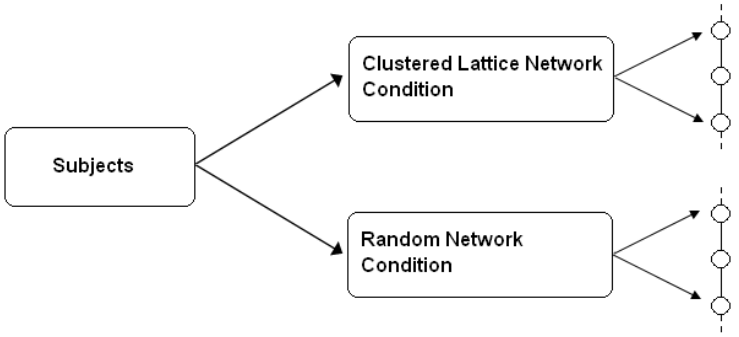

Figure S1. Schema of the experiment. Each subject is randomly assigned to a network condition, and then randomly assigned to a single node in that condition.

Subject Experience During the Experiment

The design of the experiment was unknown to the subjects. Upon arriving to the study, subjects were told that they would be matched with "health buddies" with whom they could share recommendations about on-line health resources. Subjects then provided their email addresses, agreed to a consent form, and chose usernames and avatars to represent themselves in the on-line community. Finally, each subject was shown a health buddy page,

which displayed her own user avatar and health interests, as well as the avatars and health interests of her health buddies. Figure S2 shows a health buddy page. Because of the similarity of the networks across conditions within a given trial of the study, subjects' health buddy pages showed the same number of health buddies regardless of which condition they were randomly assigned to.

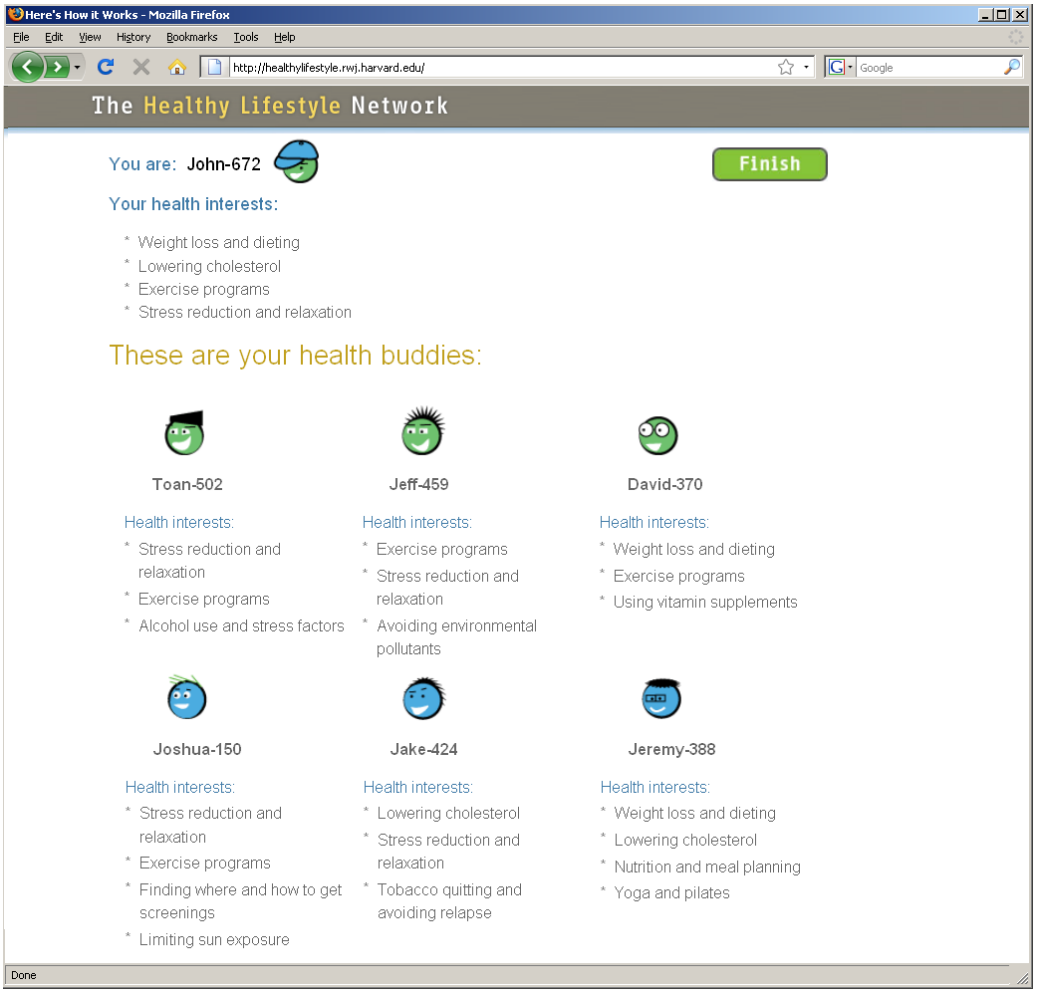

Figure S2. Screenshot of a health buddy page.

Once a subject completed the sign-up process, she received a confirmation email asking her to verify that her email address was working. This allowed me to make sure that

all subjects in the study could be reached by email. For the duration of the study, subjects only received email signals if their neighbors in the network adopted the behavior. Each subject could receive at most one email from each health buddy. Each email let the subject know which of their health buddies had registered for the health forum website, and provided a web-link to the health forum registration page. An email signal is shown in figure S3.

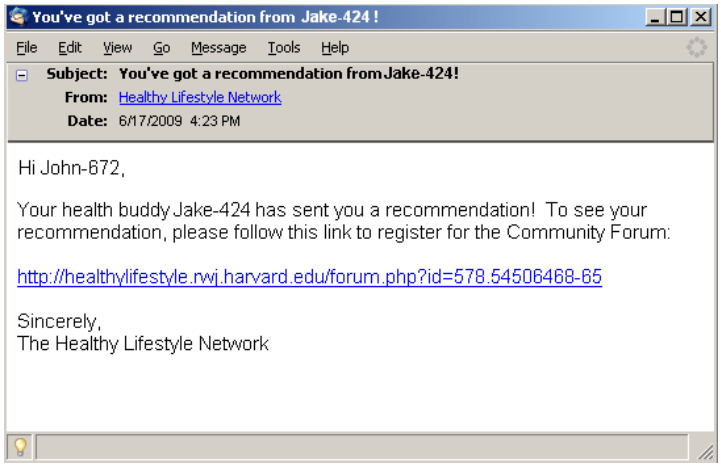

Figure S3. Email signal inviting a subject to join the health forum.

Once subjects registered for the forum, they could visit any of the websites listed under the different category headings (including “Healthy Lifestyle,” “Fitness,” “Nutrition”, “Smoking Cessation”, and “Weight Loss”), and rate the quality of the websites. Figures S4 and S5 show the registration page, and the home page, respectively, of the health forum. Every subject who registered for the forum saw the most current list of sites and ratings, available to all members. The only difference in the content of the health forum across subjects was the list of health buddies shown in the lower left panel – this was determined by which of the subject’s health buddies had already registered for the forum. The ratings for

the websites, and the listed order of the websites, changed in real-time as subjects interacted with the health forum.

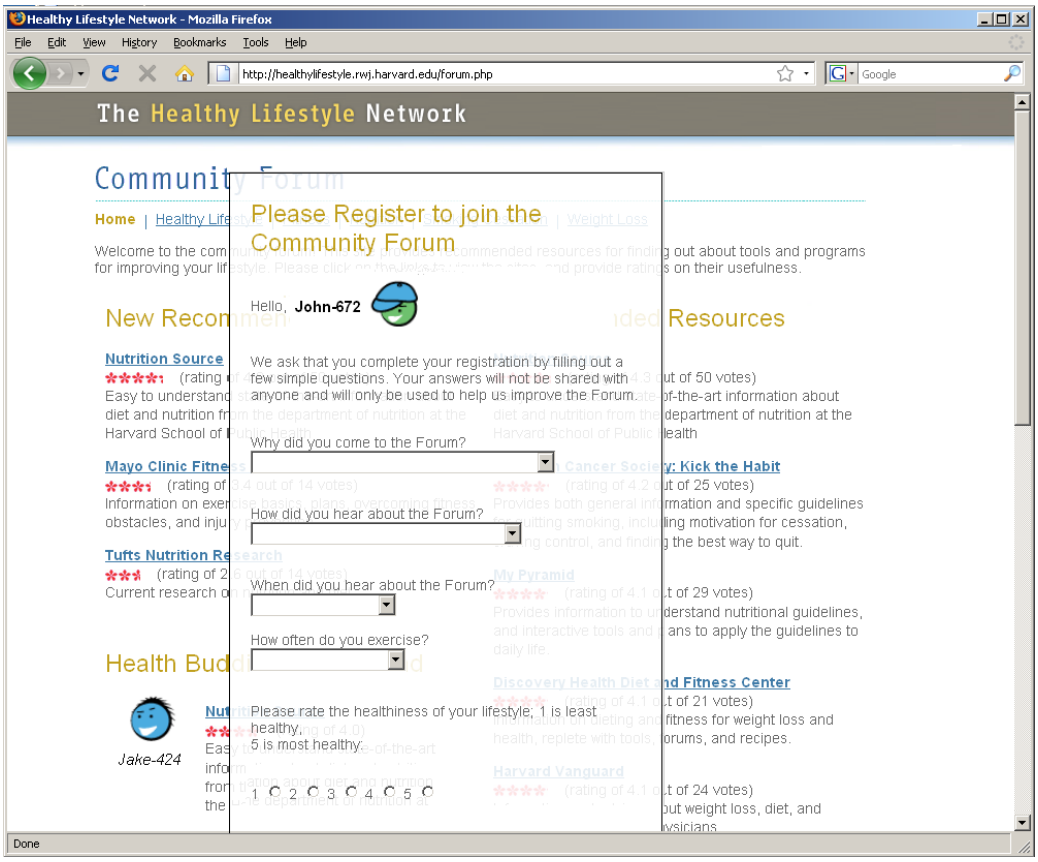

Figure S4. Registration page for the health forum website. Subjects were required to register in order to access the forum.

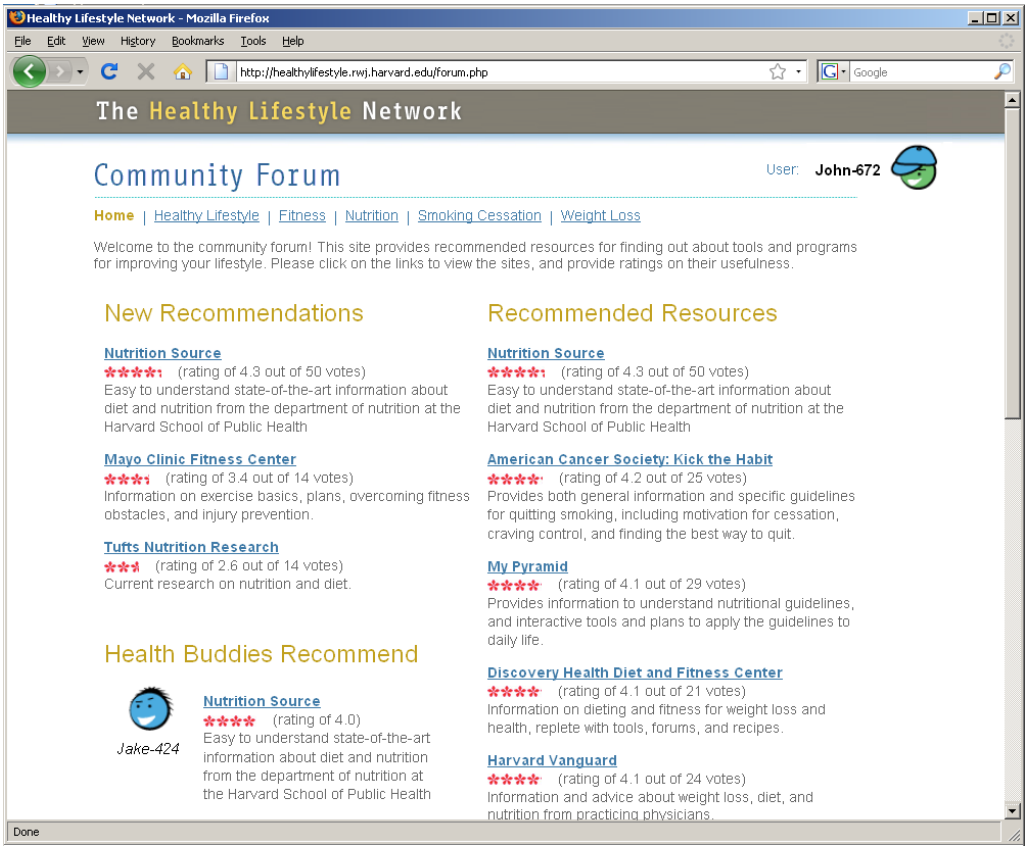

Figure S5. Home page for the health forum website.

Subjects' activity in the health forum did not affect the experiment. Once subjects registered for the forum their actions did not result in additional email signals being sent, nor could they receive additional email signals. Entry to the health forum was treated as a binary and irreversible adoption decision, which was recorded as a one time event, resulting in a single message being sent to each of the subject's health buddies who had not already adopted the behavior.

If subjects returned to the initial sign-up website for the experiment, or to the health forum website, their most recent information was displayed, and they were not given the option to re-register either for the study or for the forum. However, subjects could return to the health forum as often as they liked (which logged them in automatically once they were

registered), and were able to visit new sites, add new ratings, and see the most recent ratings by their health buddies.

Subject Recruitment

The study was run for a 124-day period from May 4, 2009 through September 5, 2009, over which time a series of recruitment campaigns were used to attract subjects to the experiment. In total, 1,528 subjects participated in the study, 764 in the clustered lattice network condition, and 764 in the random network condition. Most subjects were recruited through email advertisements sent to members of health websites such as Prevention (http://www.prevention.com), Self (http://www.self.com), Men's Health (http://www.menshealth.com), Women's Health (http://www.womenshealthmag.com), and Shape (http://www.shape.com). Figure S6 shows examples of recruitment advertisements. Additionally, a small fraction of subjects were recruited from Your Disease Risk, a cancer risk evaluation site run by Washington University in St. Louis (http://www.yourdiseaserisk.wustl.edu).

Figure S6. Advertisements used to recruit subjects to the study.

Network Structures

Each trial was comprised of a clustered lattice condition and a random network condition. Clustering in the network was measured using the clustering coefficient (*CC*, calculated using the closed triples method, *S1*), which reports the fraction of closed triples in the network. Trials 1, 2, 3, and 4 used a hexagonal lattice network (Z=6, CC=0.4) for the lattice condition (fig. S7A shows the hexagonal neighborhood structure) and a randomized version of the lattice network for the random condition (N=98, CC=0.035; N=128, CC=0.042). Randomization used the Maslov-Sneppen small world rewiring technique (*S2*,*S3*), which entails rewiring each tie to a random location in the network while preserving each node's degree, resulting in a network with the same degree distribution (Z=6), but with a completely random topology. Trials 5 and 6 used a Moore lattice network (Z=8, CC=0.43) for the lattice condition (fig. S7B shows the Moore neighborhood structure), and used the same randomization procedure described above to create the corresponding random networks (Z=8, CC=0.042). Both the hexagonal and Moore lattice networks were located on tori (i.e., toroidal surfaces), so there were no boundary effects in the networks.

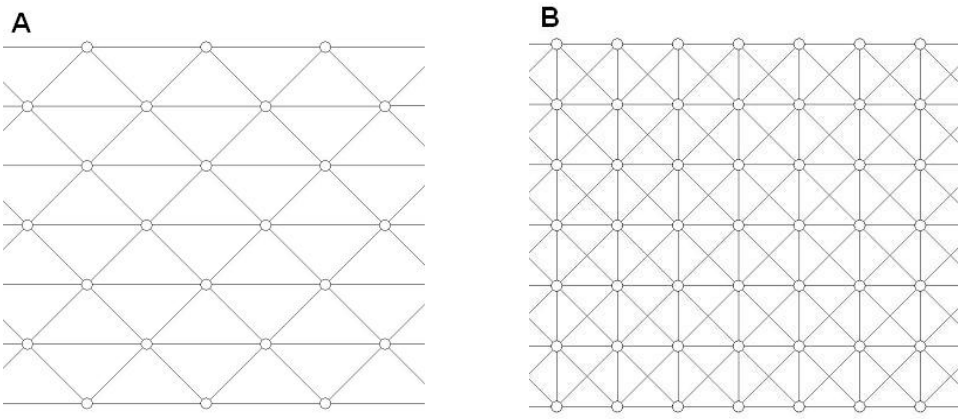

Figure S7. Neighborhood structure for the **A**) Hexagonal and **B**) Moore lattice networks.

The neighborhood structures of the Hexagonal and Moore lattices have two corresponding consequences for topology. First, as noted above, the Moore lattice has a

greater level of clustering than the Hexagonal lattice (0.43 > 0.4). Second, the maximum overlap between neighborhoods (the greatest number of neighbors shared by two non-neighbors) is 2 in the Hexagonal lattice, and 3 in the Moore lattice. This value is significant for the diffusion of complex contagions since it determines the number of redundant signals that an individual can receive from a single nearby neighborhood (*S4*,*S5*). Additional experimental studies are needed to determine how these complementary features of the lattice topologies independently affect the diffusion of behaviors.

Data Analysis

I measured the success of diffusion in terms of the fraction of the population that ultimately adopted the behavior. The fraction of adopters, S_j, in network j, is defined as:

$$S_j = \frac{\sum_{i=1}^{(N_j-1)} a_i}{(N_j - 1)}, \qquad (1)$$

where N_j is the number of nodes in network j (1 is subtracted to account for the seed node), and $a_i = 1$ iff node i adopted, otherwise $a_i = 0$. Within each trial, both networks in both conditions had the same size and degree distribution. However, across trials some networks had different sizes and degree distributions. This was done to ensure that the results from this study were not an artifact of a specific choice of neighborhood size or network size. To make comparisons across trials possible, success is measured in terms of the *fraction* of the population that adopted, not the absolute number. The success of diffusion for each condition of each trial, and mean and standard deviation of these values across all trials, are shown in Table S1.

I used the Wilcoxon Rank Sum Test (also known as the Mann-Whitney U test) to evaluate the statistical significance of differences in success across the six trials. The Wilcoxon is a non-parametric test of the likelihood that observations drawn from one population will be greater than those drawn from another population. In essence, it tests whether there is a statistical significance in the difference of the medians of two populations. Thus, it is very similar to the two-sample *t*-test, however it provides a more conservative estimate of significance since it does not rely on the assumption of normality in the distribution. The Wilcoxon test shows that the null hypothesis that there is no difference in the success of diffusion between the two conditions can be accepted with a probability of $p<0.01$.

	Random Network Condition (Total Fraction Adopted)	Clustered Lattice Network Condition (Total Fraction Adopted)
Trial 1 (N=98, Z=6)	0.3814	0.5154
Trial 2 (N=128, Z=6)	0.2677	0.3937
Trial 3 (N=128, Z=6)	0.2834	0.5275
Trial 4 (N=128, Z=6)	0.3700	0.5590
Trial 5 (N=144, Z=8)	0.5104	0.6363
Trial 6 (N=144, Z=8)	0.4825	0.5944
Mean	0.3826	0.5377
Standard Deviation	0.0995	0.0833

Table S1. Success of Diffusion in Clustered Lattice and Random Networks

To measure the rate of diffusion, I compared the time it took for the diffusion process to reach the farthest node that was reached by both conditions in a given trial. For example, in trial 1 diffusion in the random network condition reached 38.14% of the network (37 nodes), while diffusion in the clustered lattice network condition reached 51.54% of the network (50 nodes). Thus, the rates of diffusion for trial 1 are compared by evaluating the

time it took each network to reach 37 nodes. More generally, the rate of diffusion in network *j* in trial *T*, R_{Tj}, is defined as:

$$R_{T_j} = \frac{\min(S_{T_0}, S_{T_1})}{time_{T_j}[\min(S_{T_0}, S_{T_1})]}, \quad (2)$$

where S_{T_0} and S_{T_1} are the fraction of adopters in conditions 0 (random network) and 1 (clustered lattice network) of trial *T*, respectively. $time_{T_j}[\min(S_{T_0}, S_{T_1})]$ reports the time it took in network *j* of trial *T* for the behavior to reach the largest fraction of nodes reached by both networks in trial *T*. Table S2 shows the distances, times, and rates corresponding to each condition in each trial, and the mean and standard deviation of the rates of diffusion across all trials.

An alternative approach to measuring diffusion rates is to pick a specific prevalence (say 50% adoption), and compare all networks at this same prevalence point. However, since many of the random networks did not reach 50%, this comparison would omit some of the data. In order to include all of the data in such a comparison, the prevalence point would have to be below 27% adoption. Yet, since many of the networks (both clustered lattice networks and random networks) spread well past 27%, this measure does not provide a good representation of the overall diffusion processes recorded in these trials.

The approach that I adopted of comparing the network conditions on a trial-by-trial basis ensures that all networks are included in the comparison. Further, this approach also ensures that the greatest possible number of data points are included to give the most accurate picture of the rate at which the behavior spread through each of the networks. Finally, because the network conditions are already paired into trials, comparing the time it takes to reach an equivalent distance within each trial is a natural way to evaluate the rate of

diffusion across conditions. Because these rate measurements have commensurate units across trials (i.e., nodes/sec), they can be aggregated to provide a summary statistic, as shown in the last two rows of table S2.

To evaluate the significance of the differences in rates across the six trials I used the same Wilcoxon Rank Sum Test described above. The logic for using this test to compare rates is the same as it was for the evaluating the differences in success: the goal is to determine the likelihood that one condition will consistently produce observations that are greater than those for the other condition. The Wilcoxon test shows that the null hypothesis that there is no difference in the rates of diffusion between the two conditions can be accepted with a probability of $p<0.01$.

		Random Network Condition		Clustered Lattice Network Condition	
	Max. Nodes Adopted in Both Conditions	Time (seconds)	Rate (x 10^{-3} nodes/sec)	Time (seconds)	Rate (x 10^{-3} nodes/sec)
Trial 1 (N=98, Z=6)	37	1021564	0.0362	634444	0.0583
Trial 2 (N=128, Z=6)	34	1051161	0.0323	157112	0.2164
Trial 3 (N=128, Z=6)	36	1016415	0.0354	294609	0.1221
Trial 4 (N=128, Z=6)	47	386516	0.1215	152509	0.3081
Trial 5 (N=144, Z=8)	73	1803602	0.0404	101640	0.7182
Trial 6 (N=144, Z=8)	69	575523	0.1198	256616	0.2688
Mean	-	-	0.0643	-	0.2820
Standard Deviation	-	-	0.0437	-	0.2328

Table S2. Rates of Diffusion in Clustered Lattice and Random Networks

The effect of social reinforcement on the individual likelihood of adoption was calculated using the Cox proportional hazards model. The Cox model is a semi-parametric test of hazard rates, which does not assume an underlying functional form for the hazard of adoption. The baseline hazard for adoption is based on the individuals who adopted after one

signal. This hazard function was then used to evaluate the conditional hazard of adoption for individuals receiving additional signals. This test thus measures the increase in likelihood that an individual will adopt the behavior from receiving multiple social signals, conditioned on the likelihood of adoption from receiving a single social signal.

The results of the Cox Model showed that receiving a second social signal increased the probability of adoption by 1.67 times, with 95% confidence intervals ranging from 1.35 to 2.05. The null hypothesis – that there was no effect of receiving a second signal on the likelihood of adoption – can be accepted with a probability of $p<0.001$. Receiving a third signal increased the likelihood of adoption by an additional 1.32 times, with 95% confidence intervals ranging from 1.01 to 1.73. The null hypothesis can be accepted with a probability of $p < 0.05$. There was no significant effect of additional social signals on the likelihood of adoption.

To investigate the effects of social reinforcement on individuals' level of commitment to their memberships in the health forum, I compared subjects' return visits to the forum (after registering for it) based on the number of signals received from their health buddies. To do this, I grouped adopters by the number of signals they received, and compared the distribution of return visits for each group. The primary comparison was between adopters who only received 1 signal (group 1), and adopters who received multiple signals (groups 2 through 5). Each group contained all and only the adopters who received exactly that number of signals: group 2 contained adopters receiving 2 signals, group 3 contained adopters receiving 3 signals, and so on. Pairwise statistical comparisons between group 1 and groups 2 though 5 are shown in Table S3 ($p<0.01$ for all four comparisons, using the Kolmogorov-Smirnov test). These comparisons show that the null hypothesis that the results

for group 1 were drawn from the same distribution as those for groups 2 through 5 can be rejected. Table S3 also shows that all pairwise comparisons between the remaining groups were not statistically significant ($p>0.4$ for the remaining six comparisons, using the Kolmogorov-Smirnov test), indicating that the null hypothesis that the results for groups 2 for 5 were drawn from same distribution cannot be rejected.

Comparison Groups	Komolgorov-Smirnov Test *P*-value
Group 1 and Group 2	$P < 0.001$
Group 1 and Group 3	$P < 0.001$
Group 1 and Group 4	$P < 0.001$
Group 1 and Group 5	$P < 0.01$
Group 2 and Group 3	$P > 0.8$
Group 2 and Group 4	$P > 0.4$
Group 2 and Group 5	$P > 0.9$
Group 3 and Group 4	$P > 0.8$
Group 3 and Group 5	$P > 0.9$
Group 4 and Group 5	$P = 1$

Table S3. Significance of Pairwise Comparisons of Return Visits for Each Group

Additional Analyses

Additional analyses used the individual responses (i.e., "thresholds") observed in the data to estimate formal models of the collective dynamics. Because I only have a single observation for each individual, I cannot assess whether individual behavior is better captured by a deterministic threshold function, or by a more complex probabilistic response function. To approximate the collective dynamics, I created "average response functions" by aggregating the individual responses to form a collective distribution of response probabilities for each level of social stimuli (*S6*,*S7*). Different aggregation strategies

attempted include i) using the PDF of observed individual responses, ii) using proportional responses by individuals at risk for adoption at each threshold level, and iii) using aggregated probabilities based on the Kaplan-Meier survival functions. However, each of these models produced poor fits to the data. These difficulties suggest that more elaborate experimental designs are necessary, which can collect multiple observations of each individual over time, to derive better approximations of individual response functions.

Ensuring Data Quality

In all experiments researchers must take steps to ensure that the subjects do not violate the design of the experiment, either through accidental behaviors, or through malicious intent. This can be more difficult in on-line experiments, where researchers have less control over the behavior of the subjects than in traditional laboratory settings. I took several steps to ensure that the data collected were sound.

For example, I designed the health forum website so that it would be an unknown behavior to subjects before they enrolled in the experiment. However, it is possible that subjects could tell people about the health forum outside the context of the experiment. I ensured that individuals could not access the health forum without being enrolled in the experiment by giving each individual a unique log-in identifier, without which an individual could not log-in to the health forum. It is also possible that subjects could encourage peers to sign up for the study in order to get access to the health forum. To control for this possibility, subjects were asked upon registering for the health forum to report how they found out about it (from a list of options including: “From a health buddy”, “From my doctor”, “From a friend”, etc.). This allowed me to track subjects who might have come to

the study with preexisting knowledge of the health forum. Given that such knowledge would make subjects more likely to adopt after a single signal, I suspect that the presence of these individuals in the study would only have weakened the effects shown in my findings.

I also intended for social influence among participants in the study to be based entirely on the on-line networks that I created, and not confounded by peer relationships outside the study. In order to prevent people from trying to identify friends who may have also signed up for the study, or from trying to contact health buddies outside the context of the experiment, I blinded the identifiers that people used. Thus, subjects could not see the true usernames used by their health buddies. Because a message from a friend from outside of the study would be more likely to promote adoption with fewer signals, I expect that a more effective means of eliminating these relationships from the study would only increase the effects that are shown in the findings.

An additional concern with an on-line study is that subjects might return to the health forum, or to the initial study sign-up website, and their data might be re-entered into the experiment. To prevent this from happening, once a subject registered for the health forum, her unique identifier in the study was flagged, and her subsequent behavior no longer affected any aspect of the experiment. However, while subjects could no longer affect the experiment, once they joined the health forum, they were permitted to return as many times as they liked, where they were automatically logged in, and any additional recommendations and ratings they made were included in the forum to be shared with the other members.

Subjects could also return to the initial enrollment website, where they had joined the study. To ensure that no subjects were re-enrolled in the study, I placed cookies into the web-browsers of the subjects, which automatically populated the sign-up pages with their

user information, and prevented me from collecting this information a second time (even if the users altered it). Users without cookie-compatible browsers (or who had cookies disabled) were not able to sign up for the study. A subject with malicious intent could figure out how to destroy the cookies in her web-browser and then re-enroll in the study. However, I also double-checked each new user's profile against existing users (from the entire history of the study), and would not allow a new user to register if her profile information or email address matched an existing user. My attempts to detect patterns in the data that revealed the presence of this kind of malicious behavior did not find any instances of it.

Robustness to Design Choices

The six trials of this experiment represent a small portion of the parameter space of possible experiments using this design. For example, parameters like the size of the networks (N), the neighborhood size (Z), and the kind of behavior being diffused may all influence the dynamics of social diffusion. The relatively narrow range of population size (N), degree (Z) and clustering coefficient (CC) that were explored was due to significant limitations on the number of subjects that could be simultaneously recruited to participate in the study. This limited the size of the networks that could be used. Similarly, because of the need to maintain low density in the network, neighborhood size needed to remain small (within the range Z=6 to Z=8) for this study (*S4*). For larger N, there is a larger range of degree (Z) for which networks are still sparse. As discussed below, as long as networks remain sparse, increasing Z is not expected to qualitatively affect the dynamics observed in this study. Despite the limitations of this study, the results from these experiments suggest some predictions.

I anticipate that increases in the size of the network (N) will make the observed differences in diffusion between the random and clustered lattice networks more pronounced. As the networks become larger (and neighborhood size, Z, is kept constant) the probability of nodes in a random network receiving redundant signals becomes smaller, which I expect will make the effect of local structure more important for the diffusion of behaviors. Similarly, these anticipated effects are also relevant for changes to neighborhood size since I suspect that reductions in neighborhood size, Z (for a fixed population N), will equally reduce the probability that nodes in random networks will receive redundant signals. Conversely, I anticipate that decreasing network size (while keeping Z constant) will ultimately eliminate the effects of topology by making the density of the networks so great that differences in topology are insignificant. Equivalently, I expect that increasing the neighborhood size while keeping N fixed will have the same effect of increasing density to the point where topological differences are no longer significant.

More generally, for the scope condition that $N>>Z$ I expect that the results of this experiment represent the lower bound of the effects of topological structure on behavioral diffusion. As neighborhood size becomes a much smaller fraction of the total population (which naturally happens as population size increases to the order of magnitude of large societies), the confounding effects of density are eliminated. As this happens, I expect that the observed differences in the diffusion dynamics between topologically diverse network conditions will become even greater.

I suspect that the results of this experiment will also generalize to other kinds of behavior. The focus on health behavior necessitated recruitment from a population of individuals who were interested in health. However, this design could easily be replicated

for other behaviors (such as adopting a product or financial service) in which the subject pool would be drawn from a population of individuals interested in those behaviors. Further, the behavior studied in this experiment is relatively easy to do compared with health behaviors such as getting screenings, changing diet, quitting smoking, or improving insurance coverage, and is available free of charge. I expect that the more difficult, or costly, the behavior is, the more dependent the success of diffusion will be on social reinforcement from locally clustered network neighborhoods. However, this implication of the study requires additional experiments in order to be confirmed. In order to generalize the results of this study for the spread of more costly behaviors, additional empirical work needs to be done.

An important feature of the design of this experiment is that personal information about the subjects was not revealed to their health buddies. This allowed me to isolate the effects of the network topology on the dynamics of diffusion without the presence of confounding variables. But, it also raises the question of what the strength of the effects of network topology would be when allowed to interact with the effects of interpersonal relationships. New experimental designs are required to test the interaction effects of these variables (and other variables, such as gender, affect, and frequency of interaction) on the spread of social behaviors. Such experimental designs will require the ability to carefully integrate each additional variable into the study, while still isolating the effects of each mechanism on the diffusion process. I anticipate many new studies in this growing area of research.

References

S1. M. E. J. Newman, *SIAM Review* **45**, 167 (2003).

S2. S. Maslov, K. Sneppen, *Science*. **296**, 910 (2002).

S3. S. Maslov, K. Sneppen, U. Alon in S. Bornholdt, H.G. Schuster (Eds.) *Handbook of Graphs and Networks* (Wiley-VCH and Co., Weinheim, 2003).

S4. D. Centola, M. Macy, *Am. J. Sociol.* **113**, 702 (2007).

S5. D. Centola, V. Eguiluz, M. Macy, *Phys. A.* **374**, 449 (2007).

S6. P. Dodds, D. Watts. *J. Th. Bio*. **232**. 587 (2005)

S7. P. Dodds, D. Watts. *Phys. Rev Lett*. **92**. 218701 (2004)