Classification: Social Sciences / Social Sciences

Title: The Network Dynamics of Social Influence in the Wisdom of Crowds

Authors: Joshua Becker[1], Devon Brackbill[1], Damon Centola[1,2]

[1]Annenberg School for Communication, University of Pennsylvania.

[2]School of Engineering, University of Pennsylvania.

Corresponding Author: Damon Centola

dcentola@asc.upenn.edu

3620 Walnut Street, Suite 200, Philadelphia, PA 19104

(215) 898-7954

Keywords:

collective intelligence, wisdom of crowds, social networks, collective behavior

Abstract

A longstanding problem in the social, biological, and computational sciences is to determine how groups of distributed individuals can form accurate judgments. Since Galton's discovery of the "wisdom of crowds" [Galton F (1907) Nature 75:7] theories of collective intelligence have suggested that in order for a group to be accurate, individuals within the group must be either independent, with uncorrelated beliefs, or diverse, with negatively correlated beliefs [Page, S (2007) *The Difference*]. Previous experimental studies argued that social influence undermines the wisdom of crowds, showing that individual estimates became more similar when subjects observed each other's beliefs, reducing diversity without a corresponding increase in group accuracy [Lorenz J et al (2010) Proc Natl Acad Sci 108:22]. In contrast, we find general network conditions under which social influence can improve group estimates, even as individual opinions become more similar. We present theoretical predictions and large scale experimental results showing that in decentralized communication networks, group estimates become more accurate as a result of exposure to social information. We also present results showing that in centralized networks, the influence of central individuals dominates the collective estimation process, and group estimates are as likely to increase in error as they are to become more accurate.

Significance Statement

Since the discovery of the wisdom of crowds over 100 years ago, theories of collective intelligence have held that group accuracy requires either statistical independence or informational diversity among individual beliefs. Empirical evidence suggests that allowing people to observe the beliefs of others leads to increased similarity of individual estimates, reducing independence and diversity without a corresponding increase in group accuracy. As a result, social influence is expected to undermine the wisdom of crowds. We present theoretical predictions and experimental findings demonstrating that in decentralized networks, social influence generates learning dynamics that reliably improve the wisdom of crowds. We identify general conditions under which influence, not independence, produces the most accurate group judgments.

Main Text

Since Galton's discovery of the "wisdom of crowds" over 100 years ago (1), results on crowdsourcing (2,3), prediction markets (4), and financial forecasting (5,6) have shown that the aggregated judgment of many individuals can be more accurate than the judgments of individual experts (1,7-9). Statistical explanations for this phenomenon argue that group accuracy relies on estimates taken from groups where individuals' errors are either uncorrelated or negatively correlated, thereby preserving the diversity of opinions in a population (10). Thus, while individuals may have estimates both far above and far below the true value, in aggregate these errors cancel out, leaving an accurate group judgment (2,10,11).

Recent experimental evidence has suggested that the wisdom of crowds may be undermined by processes of social influence, in which people exchange information about their estimates and revise their judgments to align with one another (12-14). When social influence leads to correlated errors, both independence and diversity are reduced, which has been argued to compromise the reliability of the group judgment (10,12-19). In direct contrast with these results, however, theoretical models of social learning (20-22) have suggested that the effects of social influence on collective decisions vary based on the structure of the interaction network, predicting that under the right conditions, social learning can lead a group's median judgment to improve (21-25).

This prediction derives from the assumption that when people learn about the beliefs of others, they revise their own beliefs to become more similar to their social referents (12,13,26,27). Following the DeGroot model of social learning, this theory suggests that each individual's revisions are based on a weighted average of their own belief and the beliefs of their social referents (20). Thus, an individual's revision is determined in part by the amount of weight they place on their own belief relative to social information. When this "self-weight" is independently and identically distributed throughout a population, and the population is embedded in a decentralized social network (i.e., one in which everyone is equally connected), this model predicts that belief distributions will converge on the statistical mean of the initial, independent beliefs (SI Appendix). Thus, if the initial group mean is accurate, exposure to social influence will lead individuals in the group to become more accurate, improving the accuracy of the group's median even as the group mean remains fixed (SI Appendix).

We build on the DeGroot model to generate novel theoretical predictions for how social influence can affect the accuracy of group judgments. We show that if individuals' self-weights are not identically and independently distributed in the population, but are instead correlated with individual accuracy, then social influence may not only improve the median judgment by bringing it toward the mean, but may also result in the mean of the population estimate becoming more accurate. This prediction builds on DeMarzo, Vayanos, and Zweibel's (21) notion of "social influence weight," which identifies the amount of influence that each individual in a network has on the collective belief. Because self-weight contributes to social influence weight, a correlation between accuracy and self-weight means that more accurate individuals will be more influential in the group estimate, resulting in a direct improvement in the accuracy of the group mean.

Our predictions also show that this process can go awry if some individuals in a population are more prominent than others, giving them disproportionate levels of social influence in the population. Theoretical results suggest that when networks are highly "centralized" in this way, instead of efficiently aggregating all available information, populations are biased toward the beliefs of the central individuals (21), which can significantly influence the accuracy of the collective judgment (SI Appendix). This effect of centralization on group estimates has been predicted by a variety of social learning models, including both fixed (21,24) and growing (22,25) networks, as well as models of both discrete choice (24,25) and continuous estimation (21,22).

We test these theoretical predictions using a Web-based experimental design. We recruited subjects to participate in a series of large group estimation tasks, and compared the effects of social influence in both centralized and decentralized networks to a control condition in which there was no social influence. Consistent with previous work, our theoretical results predict that centralized networks will exhibit a bias toward the beliefs of central individuals. However, in contrast to prior work showing that social influence undermines group accuracy (12-19), we predict that social influence in decentralized networks will improve the accuracy of the group median (SI Appendix). Moreover, we also predict that social influence can produce systematic improvements in the accuracy of the group mean if the individual revision process is not identically and independently distributed, but is correlated with individual accuracy. As described below, our experimental design permits a direct test of these theoretical predictions based on our extensions of the DeGroot model.

Theoretical Model

We build on DeGroot's (20) formalization of local information aggregation, in which subject *i* updates their response estimate, $R_{t,i}$, after being exposed to the estimates of their network neighbors, $\overline{R}_{t,j \in N_i}$ We define a subject's revision process with three components: their own estimate; the estimates of network neighbors; and "self-weight," or the amount of weight they place on their own estimate relative to the estimates of their network neighbors. In this model, each subject responds to social information by adopting a weighted mean of their own estimate and the estimates of their neighbors, according to the rule:

$$R_{t+1,i} = \alpha_i \times R_{t,i} + (1 - \alpha_i) \times \overline{R}_{t,j \in N_i},$$

where the value $R_{t,i}$ indicates the response of subject *i* at time *t*; α_i indicates the self-weight a subject places on their own initial estimate; (1- α_i) indicates the weight they place on the average estimate of their network neighbors; and $\overline{R}_{t,j\in N_i}$ indicates the average estimate of subject i's network neighbors at time t. Outcomes are therefore determined by three parameters: the communication network (i.e., who can observe whom), the distribution of initial estimates, R_1, and the distribution of self-weights, α_i.

At the population level, this model states that the group estimate after t revisions can be calculated as a function of a weighted, directed network of social influence (20). In this social influence network, a tie exists from node i to node j if i can observe j in the communication network. The tie has a numeric value that indicates the weight that node i places on the estimate of node j which is determined by α_i. For any given node i, the sum of the outgoing tie weights equals $(1 – \alpha_i)$. Consistent with previous implementations of this model (20-22), we represent the self-weight that node i places on its own estimate (α_i) as a "self-tie" from i to i. The set of each node's outgoing tie weights (including their self-tie) sums to 1. The sum of each node's incoming ties (including their self-tie) is proportional to their overall influence in the network during each round of revisions – i.e., their "social influence weight" – which is defined as each subject's influence in the collective estimation process (21). Because this sum includes the subject's self-weight, each subject's influence in the collective estimation process is determined in part by how heavily they weight their own opinion compared to the social information they receive.

This concept of social influence weight comes from the properties of the DeGroot model, where members of a population revise their estimates indefinitely according to the process above. Through this revision process, the DeGroot model predicts that in a wide range of network structures all members of the population will asymptotically converge on a single shared estimate (20). The collective estimate after social influence is a weighted mean of the initial independent estimates (21). Each individual's social influence weight is defined by the size of the contribution that their initial (independent) estimate makes to the final collective estimate (21). The relationship between self-weight and social influence weight reflects the fact that when a subject places more weight on their own individual belief, they adjust their belief less in response to others, and thereby contribute more weight to the group estimate (21).

In decentralized networks – defined as networks where everyone has the same number of ties (28) – the properties of the model described above indicate that the arithmetic mean of a group's estimate distribution will remain unchanged even as social influence leads individuals' estimates to become more similar. This prediction (convergence toward the mean) holds under the assumption ~~either~~ that self-weight is identically and independently distributed (i.i.d.) throughout a population (SI Appendix). If this process accurately characterizes the effects of influence on the wisdom of crowds, and the initial group mean is accurate, then social influence in decentralized networks allows individuals to increase the accuracy of their estimates without any deleterious effects on group-level accuracy. One consequence of this is that the median of the group estimate can improve while the mean stays fixed (SI Appendix).

We extend these predictions by analyzing what happens when this i.i.d. assumption is violated – i.e., there is non-i.i.d. heterogeneity in the degree to which individuals revise their estimates based on the estimates of others. Our results predict that if an individual's self-weight is correlated with their accuracy, social influence dynamics may not only be able to improve the median judgment by bringing it toward the mean, but may also result in the mean of the population estimate becoming more accurate as a function of social influence (SI Appendix).

Experimental Design

We recruited 1,360 participants from the World Wide Web to take part in a series of estimation challenges. Subjects were randomized either to one of two experimental social network conditions, or to a control condition. In all conditions, participants were prompted to complete estimation tasks and were awarded a monetary prize based on the accuracy of their final estimate. In the network conditions, participants were placed into either a decentralized network, in which everyone had equal connectivity, or a centralized network, in which a highly connected central member had a disproportionate number of connections (see Materials and Methods and Figs. S1 and S7).

Each social network contained 40 subjects. Within each network, all subjects were simultaneously shown the same image prompt (e.g., a plate of food) and asked to estimate a numerical quantity (e.g., the caloric content) (SI Appendix). There were three rounds for each estimation task. In Round One, participants provided an independent estimate based on the prompt. In both network conditions, participants were then shown the average estimate of the peers directly connected to them in their social network, and prompted to submit their answers again in Round Two. Subjects were then shown the average of their peers' revised estimates, and prompted to submit a final estimate in Round Three. Thus, for each question, participants provided one independent estimate and two estimates after exposure to social information, for a total of three estimates per question. Subjects were not provided with any information about their social networks, which ensured that the subject experience was identical across the two network conditions.

Subjects who were randomized to the control condition were not placed into social networks, but were instead given the opportunity to answer the same questions without being exposed to social influence. These participants were still given the opportunity to revise their initial answer two times, providing a total of three independent estimates per question. All control participants observed the same sets of questions in the same order as participants embedded within the experimental networks. More generally, the subject experience in the control condition was identical to that of subjects in other conditions, except that participants were not given any social information.

Each experimental trial of the study consisted of an identical set of questions provided to one decentralized network (40 individuals) and one centralized network (40 individuals). For each set of questions that was asked in the experimental trials, we also collected responses from 40 independent individuals in the control condition, who collectively formed a “control group” for that set of questions.

Because subjects in the network conditions were not statistically independent, all analyses of collective estimates in the network conditions were conducted at the group level (SI Appendix). Moreover, because each network completed multiple estimation tasks within an experimental trial, we cluster our main analysis at the trial level such that each network provides a single, independent observation (SI Appendix). In total, we conducted 13 experimental trials, comprising 520 subjects in each network condition (1040 experimental subjects in total). In 6 of the experimental trials, subjects answered 4 questions in each trial, where each question set was unique. In the remaining 7 trials, subjects answered 5 questions in each trial, using 2 unique questions over repeated trials. In total, this produced 8 unique question sets.

Control groups were conducted corresponding to each unique question set, producing 8 control groups, each of size 40 (320 control subjects in total) (See Methods and Materials). Because subjects in the control groups were independent from each other, fewer overall subjects were required for the control analyses (see Methods and Materials; SI Appendix). Nevertheless, for proper comparison with the experimental conditions, we still conducted our control trials with subjects in groups of 40, and conduct our analyses at the group level (SI Appendix).

We measure the cumulative effect of social influence on collective judgments by comparing the initial estimates of each group (i.e., in Round One of our study), with the final estimates of each group after two rounds of revision (i.e., in Round Three). For results where we report percent change, all comparisons are made between final estimates (i.e., Round Three) and independent estimates (i.e., Round One), so that percent change is measured as the magnitude of the change in the estimate divided by the magnitude of the initial estimate (SI Appendix). To facilitate comparisons across different estimation tasks of different scales (i.e., some questions have true answers over 1000, while some questions have true answers under 100) we normalize all estimates, dividing them by the standard deviation of the independent responses for each question (SI Appendix). All reported changes in error are therefore measured in terms of the distance of each estimate from the truth, represented as the number of standard deviations (s.d.) away from the true answer (comparable to a z-score).

Results

Social network structure significantly affected the wisdom of crowds. We found both that decentralized networks showed the predicted increase in collective accuracy, and that centralized networks exhibited the predicted bias toward the beliefs of central individuals. We begin our analysis by confirming that in the independent round (i.e., Round One of all trials) groups exhibited the wisdom of crowds. Consistent with earlier studies (1,6-9,14), we found that, on average, both the mean and the median of each group's estimate was more accurate than the majority of its members (SI Appendix). In the results that follow, we analyze how social influence affected the trajectory of group estimates in each of the network conditions.

Decentralized Networks. Social influence dramatically reduced the diversity of group estimates. As shown in Fig. 1D, two rounds of revision significantly narrowed the standard deviation of responses (N=13 trials, P<0.001, Wilcoxon signed rank test), producing a 43% reduction in the average standard deviation between Round One and Round Three. This sizable reduction in diversity replicates the main finding from previous experimental research on social influence in the wisdom of crowds (14).

However, this reduction in diversity did not undermine the wisdom of crowds. Rather, consistent with previous research (29,30), we found that social influence in decentralized networks produced significant improvements in individual accuracy. Across all 13 trials with decentralized networks, average individual error was significantly lower in Round Three than it was in Round One, decreasing by 23% on average (N=13 trials, P<0.001, Wilcoxon signed rank test). In addition to these individual level improvements, we also found that the average error of each group's median estimate was significantly lower in Round Three (0.67 s.d.) than in Round One (0.76 s.d.) (*N*=13 trials, *P*<0.001, Wilcoxon signed rank test), resulting in a 12% decrease in average error, as shown in Figs. 1A and 1C.

In our analysis of how social influence produced these group-level improvements in the median, our initial expectation was that self-weights were independently and identically distributed within each network. On this assumption, the DeGroot (20) model predicts that social influence in decentralized networks can improve the group median by pushing it towards the mean of the group's independent estimate, which is not expected to change (SI Appendix). Remarkably, however, we found that, on average, each group's mean estimate also became more accurate. After two rounds of exposure to social influence, the average error of the group mean at Round Three (0.62 s.d.) was significantly lower than at Round One (0.69 s.d.) (*N*=13 trials, *P*<0.01, Wilcoxon signed rank test), resulting in a 10% reduction in the average error of the group mean. These findings can be explained with the DeGroot model by observing that individuals' self-weights were not identically distributed in the population.

Figure 2 shows that across all network conditions the magnitude of an individual's revisions from Round One to Round Three was significantly correlated with the magnitude of their initial error (*N*=4340 estimates by 1040 subjects, ρ=0.41, 95% CI [0.39, 0.43], *P*<0.001, Analysis of Covariance). Because each individual completed multiple estimation tasks, we measure this relationship between individual accuracy and revision magnitude after controlling for correlation between estimates by the same individual (SI Appendix). The results (Fig.2) show that initially accurate individuals made smaller revisions to their estimates, while initially inaccurate individuals made larger revisions. Consistent with the

DeGroot model, one explanation for this revision pattern is that individuals who were more accurate had greater self-weight in their revisions than individuals who were less accurate. This explanation is consistent with the observed behavior, however our analysis also needs to account for the observation that individuals who were more accurate also had estimates that were closer to their observed neighborhood average. Consequently, the positive correlation between error and revision magnitude may be due to the fact that subjects whose initial estimates were farther from their neighborhood average were inclined to make larger revisions, rather than to the fact that more accurate individuals had a stronger self-weighting.

To control for this potentially confounding effect, we measured the partial correlation between error and revision magnitude, while holding constant the distance between the subject's initial estimate and the initial neighborhood estimate. The inset in Fig. 2 shows that even with this statistical control, more accurate individuals still made smaller revisions to their estimates than less accurate individuals (N=4340 estimates by 1040 subjects, ρ=0.25, 95% CI [0.22, 0.28], P<0.001, Analysis of Covariance). This suggests that accurate individuals placed more weight on their own estimates and less weight on social information (SI Appendix). By contrast, less inaccurate individuals had a lower self-weight, and were more influenced by social information. For clarity, we refer to this partial correlation between accuracy and self-weight as the "revision coefficient."

As discussed above, each individual's social influence weight in the network is determined in part by their self-weight, so that individuals who place more weight on their own estimate are also more influential in the collective estimate. When considered in the context of our theoretical model, the correlation shown in Fig. 2 indicates that more accurate individuals had a larger social influence weight in the network, which can pull the group estimate toward a more accurate mean (SI Appendix). These analyses suggest a direct positive relationship between the average revision coefficient among the members of a group and the expected improvement in the accuracy of the group mean. Fig. 3A shows, for decentralized networks, the correlation between the improvement in the group mean for each question, and the group's revision coefficient for that question, for each of the 59 group estimation tasks completed in decentralized networks. Because each group completed multiple estimation tasks, these analyses control for correlations across multiple estimates made by the same group (See SI Appendix).

Consistent with our theoretical expectations, the correlation shown in figure 3A indicates that in decentralized networks, groups with higher revision coefficients also exhibited larger improvements in group accuracy (N=59 estimation tasks, ρ=-0.71, 95% CI [-0.82, -0.56]. By contrast, figure 3B shows that centralized networks (as discussed below) exhibited no significant correlation between a group's average revision coefficient and a change in group accuracy (N=57 estimation tasks, ρ=-0.16, 95% CI [-0.33, 0.10]).

Figure 3A indicates that, in decentralized networks, the greater the correlation between individual accuracy and self-weight, the more likely it is that the group mean will improve. Additional simulation analyses, which are provided in the SI Appendix (Fig. S9), show that in decentralized networks a positive revision coefficient is sufficient to produce increases in group accuracy consistent with our empirical findings. Notably, across all experimental trials, the average revision coefficient for all subjects was positive (SI Appendix) suggesting that in very large populations with decentralized networks, social influence is likely to generate consistent improvements in the accuracy of the group mean.

Control Condition. These improvements in both the mean and the median, as well as the accuracy of individuals' estimates, all contrast with the results from the control condition (i.e., without social influence). Subjects in the control condition were able to revise their answers several times, but were not provided any information about the estimates of other participants. Between Round One and Round Three, groups in the control condition showed only a small (3%) decrease in average standard deviation (SI Appendix), which was significantly smaller than the reduction in diversity in decentralized networks (43%) and centralized networks (42%) (N=21, 13 experimental and 8 control trials, P<0.001 for both comparisons, Wilcoxon rank sum test). The opportunity for revision produced a small (3%) decrease in average individual error even in the absence of social information (N=8 control trials, P<0.001, Wilcoxon signed rank test). However, this improvement was significantly smaller than the 23% improvement by individuals in decentralized networks (N=21, 13 experimental and 8 control trials, P<0.001, Wilcoxon rank sum test). Moreover, in the control condition, these individual improvements produced no significant changes in the accuracy of either the group mean (P>0.94) or the group median (P>0.64) (complementary analyses provided in the SI Appendix). These results indicate that the improvements in collective judgment observed in decentralized networks are not explained by independent learning effects, but are due to the network dynamics of social influence.

Centralized Networks. In each centralized network, one randomly selected participant was given disproportionate exposure to the rest of the network by being given many more network ties than other subjects (see SI Appendix). Because these central individuals had more network ties than other individuals, they had much greater weight in the resulting network of social influence. As expected, the diversity of estimates in centralized networks (shown in Fig. 1D) significantly decreased after social influence (N=13 trials, P<0.001, Wilcoxon signed rank test), reducing the average standard deviation by 42%. However, averaged over all trials, social influence in centralized networks did not reliably improve either the group mean (P>0.63) or the group median (P>0.78). Instead, as predicted by the DeGroot model, the effects of social influence were determined by the initial estimates of the central individual.

To analyze these effects, we divided the group estimates in centralized networks into two categories, based on the initial estimate of the central nodes. In one category ("center toward truth") the influence of the central node is expected to increase the accuracy of the group mean. This category includes estimates in which the central node was more accurate than the group mean, and also estimates in which the central node was less accurate, but was on the opposite side of the truth from the group mean. For instance, if the true value is 100 and the group mean is 90, a central node with an estimate of either 105 (more accurate) or 120 (less accurate) will pull the group toward the truth (see SI Appendix, Fig S8). The second category ("center away from truth") includes trials in which the estimate of the central node pulled the group mean away from the truth. For instance, if the estimate of the central node is instead 70. This analytical strategy was used to identify the effects of social influence on both the group mean and the group median, as reported below.

All 13 trials produced responses to at least one question in which the central individual was away from truth relative to the group estimate, while only 12 trials produced responses where the central individual was towards truth. Accordingly, our analyses for each category use N=13 trials and N=12 trials, respectively. As shown in Fig. 1B and 1C, when the central individual's estimate was toward truth, the average error of the group mean after social influence (0.32 s.d.) was 43% lower than the average error of the group mean before social influence (0.56 s.d.), producing a significant increase in group accuracy (N=12 trials, P<0.01, Wilcoxon signed rank test). Correspondingly, the same analysis for the median showed that the error of the median also decreased significantly by 48% in these group estimations from Round One (0.70 s.d.) to Round Three (0.36 s.d.) (N=12 trials, P<0.01, Wilcoxon signed rank test). Similarly, when the central individual provided an estimate that was away from truth, social influence increased the error of the group mean by 19% and the error of the median by 32% (Fig. 1B and 1C), significantly reducing the accuracy of both the mean and the median of estimates (N=13 trials, P<0.01 for both comparisons, Wilcoxon signed rank test).

Figure 3C shows the effects of the central node on the collective estimate for each of the 57 estimation tasks in which the central node offered a response (SI). As above, because each group completed multiple estimation tasks, these analyses control for correlations between multiple estimations made by the same group (SI Appendix). The positive slope in Fig. 3C (N=57 estimation tasks, ρ=0.92, 95% [0.88, 0.95]) indicates that the group estimates in centralized networks moved toward the initial belief of the central individual – i.e., higher estimates by the central node made the group mean increase, while lower estimates made the group mean decrease.

Robustness. To conclude our analyses, we examined the robustness of our theoretical and experimental findings under variations in the network parameters, such as average degree, graph density, and population size. Graph density and average degree had no effect on the results (see Figs S11 and S12). However, we found that the effects of social influence on the wisdom of crowds are significantly strengthened with larger population sizes (see SI Appendix). Our analyses indicate that recent small group studies arguing that social influence undermines the wisdom of crowds (even in a decentralized network) (14) were insufficiently statistically powered to identify the improvements in collective accuracy that we found (see Fig. S13). Additional simulation analyses as well as supplementary analyses of the publicly available data from these studies (SI Appendix) show that these effects of population size can both explain the negative findings from previous experiments using small groups, and demonstrate the generalization of our positive results to larger population sizes.

Discussion

Our study differs in several respects from previous work on the network dynamics of collective intelligence. Unlike research on social coordination (31-33) and group problem solving (34-36), our study does not consider situations where social interaction is necessary for groups to achieve a collective outcome. Instead, we identify how the network dynamics of social influence can affect collective estimation tasks in situations where social influence has been predicted to have a negative effect on the quality of group judgments (2,12-19). Our finding that groups have the ability to generate accurate estimates even in the presence of social influence has useful implications for the design of several kinds of collective decision processes. As described in previous studies (14), if social influence did indeed undermine the wisdom of crowds, then democratic institutions and organizational decision procedures could be improved by preventing people from communicating during a voting process (14). Based on these ideas, commercial and non-profit organizations have implemented automated aggregation tools in order to collect individuals' independent beliefs in ways that minimize the information exchanged between them (37). Our findings argue against this approach to aggregation. In contrast, we have shown how social learning in networks can amplify the influence of accurate individuals, leading to both individual and collective judgments that are more accurate than those which could typically be obtained by independent aggregation alone. We therefore anticipate that process interventions within political discussion settings (38) and organizational decision contexts (2,39) may benefit more from approaches that manage communication networks, rather than approaches that attempt to increase independence in the aggregation process.

Materials and Methods

All subjects who participated in this study provided informed consent during the registration process, and all procedures in this study were approved by the Institutional Review Board of the University of Pennsylvania. Upon entering the experimental platform, participants were randomly assigned to one of three conditions – a decentralized random network, a centralized network, or a control condition (SI Appendix). Once placed into a condition, players interacted in real time for a period of approximately 15 minutes. For each question, participants first provided an independent estimate without any social information. In the network conditions, participants observed the average response of the peers immediately connected to them in a social network and were prompted to submit their answers again. Subjects were exposed to two rounds of social influence before they submitted their final answer, providing a total of three responses to each question. In the control condition, participants were given three opportunities to respond, but were not provided any social information. Monetary rewards were based on the accuracy of subjects' final response to each question.

To ensure that our findings are robust to variations in the distribution of estimates, we conducted two sets of experimental trials, using questions that generate distributions with different shapes. In the first set of trials, subjects were given count-based questions (e.g., "how many candies are in this jar?"). Because these are zero-bounded on the left and unbounded on the right, count-based questions generate highly skewed distributions (1,12), in which the median is able to improve even if the mean remains unchanged (SI Appendix). In the second set of trials, we asked participants to provide responses to percentage based question (e.g., "what percentage of people in this photograph are wearing hats"). These responses are constrained to fall between zero and one hundred, and did not produce any systematic skew in the distribution of estimates (SI Appendix).

A single experimental trial consisted of 40 individuals placed into a decentralized network and 40 individual placed into a centralized network, all of whom were given the same question set. A control group consisted of 40 independent individuals who were all given the same question set as the 80 subjects in the corresponding experimental trial. Since the subjects in a control trial were independent from one another, only one control trial was conducted for each question set.

In trials where we provided count-based estimation tasks, each group completed four tasks. We conducted 6 independent experimental trials of this kind of task, with four questions each, producing a total of 24 count-based estimations by decentralized networks, and 24 count-based estimations by centralized networks. We used a unique question set for each trial, yielding 6 unique question sets. To create independent control groups for each question set, we ran 6 independent control groups, each with 40 individuals, producing 24 control group estimations.

In trials where we used percentage-based estimation tasks, each group completed five estimation tasks. We conducted 7 independent experimental trials of this kind of task, with five questions each, producing a total of 35 percentage-based estimations by decentralized networks, and 35 percentage-based estimations by centralized networks. We used 2 unique question sets, which were randomly assigned across trials. One set was used in three of the trials, the other was used in four of the trials. To create independent control groups for each question set, we ran 2 independent control groups, each with 40 individuals, producing 10 control group estimations. Because control groups are composed of statistically independent individuals, we only require a single control group for each question set to compare to the experimentally replicated trials. In total, we observed 59 estimations by decentralized networks, 59 estimations by centralized networks, and 34 estimations by control groups.

Supporting Information is available as an appendix to this document.

Acknowledgments: We thank D. Helbing, F. Schweitzer and A. van de Rijt for comments, A. Wagner and R. Overbey for development assistance, and two anonymous reviewers for valuable suggestions that improved this article.

References and Notes:

1. Galton F (1907) Vox populi. *Nature* 75(7):450–51.

2. Sunstein CR (2006) *Infotopia: How many minds produce knowledge.* (Oxford University Press).

3. Pentland A (2014) *Social physics: How good ideas spread-the lessons from a new science.* (Penguin).

4. Wolfers J, Zitzewitz E (2004) Prediction markets. *J Econ Perspect* 18(2):107–126.

5. Kelley EK, Tetlock PC (2013) How wise are crowds? Insights from retail orders and stock returns. *J Finance* 68(3):1229–1265.

6. Nofer M, Hinz O (2014) Are crowds on the internet wiser than experts? The case of a stock prediction community. *Journal of Business Economics* 84(3):303–338.

7. Sjöberg L (2009) Are all crowds equally wise? A comparison of political election forecasts by experts and the public. *Journal of Forecasting* 28(1):1–18.

8. Herzog SM, Hertwig R (2011) The wisdom of ignorant crowds: Predicting sport outcomes by mere recognition. *Judgment and Decision Making* 6(1):58.

9. Mellers B, et al. (2014) Psychological strategies for winning a geopolitical forecasting tournament. *Psychological Science* 25(5):1106–1115.

10. Page SE (2008) *The difference: How the power of diversity creates better groups, firms, schools, and societies.* (Princeton University Press).

11. Hong L, Page SE (2008) Some microfoundations of collective wisdom. *Collective Wisdom*, eds. Landemore H, Elster J. (Cambridge University Press), pp. 56–71.

12. Jenness, A. (1932). The role of discussion in changing opinion regarding a matter of fact. Journal of Abnormal and Social Psychology, 27(3), 279-296.

13. Myers DG, Bishop GD (1971) Enhancement of dominant attitudes in group discussion. *J Pers Soc Psychol* 20(3):386.

14. Lorenz J, Rauhut H, Schweitzer F, Helbing D (2011) How social influence can undermine the wisdom of crowd effect. *Proc Natl Acad Sci USA* 108(22):9020–9025.

15. Janis IL (1982) *Groupthink: Psychological studies of policy decisions and fiascoes.* (Houghton Mifflin Boston) Vol. 349.

16. Schkade D, Sunstein CR, Kahneman D (2000) Deliberating about dollars: The severity shift. *Columbia Law Rev* pp. 1139–1175.

17. Sunstein CR (2002) The law of group polarization. *J Polit Philos* 10(2):175–195.

18. Baddeley M (2010) Herding, social influence and economic decision-making: sociopsychological and neuroscientific analyses. *Philos Trans R Soc Lond B Biol Sci* 365(1538):281–290.

19. Moussaïd M, Kämmer JE, Analytis PP, Neth H (2013) Social influence and the collective dynamics of opinion formation. *PLoS One* 8(11):e78433.

20. DeGroot MH (1974) Reaching a consensus. *J Am Stat Assoc* 69(345):118–121.

21. DeMarzo PM, Vayanos D, Zwiebel J (2003) Persuasion Bias, Social Influence, and Unidimensional Opinions. *Q J Econ* 118(3):909–968.

22. Golub B, Jackson MO (2010) Naive learning in social networks and the wisdom of crowds. *Am Econ J Microecon* 2(1):112–149.

23. Bala V, Goyal S (1998) Learning from neighbours. *Rev Econ Stud* 65(3):595–621.

24. Mossel E, Sly A, Tamuz O (2015) Strategic learning and the topology of social networks. *Econometrica* 83(5):1755–1794.

25. Acemoglu D, Dahleh MA, Lobel I, Ozdaglar A (2011) Bayesian learning in social networks. *Rev Econ Stud* 78(4):1201–1236.

26. Deutsch M, Gerard HB (1955) A study of normative and informational social influences upon individual judgment. *J Abnorm Soc Psychol* 51(3):629.

27. Price V, Nir L, Cappella JN (2006) Normative and informational influences in online political discussions. *Commun Theory* 16(1):47–74.

28. Freeman LC (1978) Centrality in social networks conceptual clarification. *Soc Networks* 1(3):215–239.

29. Farrell S (2011) Social influence benefits the wisdom of individuals in the crowd. *Proc Natl Acad Sci USA* 108(36):E625–E625.

30. Gürçay B, Mellers BA, Baron J (2015) The power of social influence on estimation accuracy. *J Behav Decis Mak* 28(3):250–261.

31. Judd S, Kearns M, Vorobeychik Y (2010) Behavioral dynamics and influence in networked coloring and consensus. *Proc Natl Acad Sci USA* 107(34):14978–14982.

32. Dall'Asta L, Baronchelli A, Barrat A, Loreto V (2006) Nonequilibrium dynamics of language games on complex networks. *Phys Rev E* 74(3):036105.

33. Centola D, Baronchelli A (2015) The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proc Natl Acad Sci USA* 112(7):1989–1994.

34. Bavelas A (1950) Communication patterns in task-oriented groups. *J Acoust Soc Am* 22(6):725-730.

35. Lazer D, Friedman A (2007) The network structure of exploration and exploitation. *Adm Sci Q* 52(4):667–694.

36. Shore J, Bernstein E, Lazer D (2015) Facts and figuring: An experimental investigation of network structure and performance in information and solution spaces. *Organization Science* 26(5):1432–1446.

37. Bonabeau E (2009) Decisions 2.0: The power of collective intelligence. *Sloan Manage Rev* 50(2):45.

38. Fishkin JS, Luskin RC (2005) Experimenting with a democratic ideal: Deliberative polling and public opinion. *Acta Politica* 40(3):284–298.

39. Green K, Armstrong J, Graefe A (2007) Methods to elicit forecasts from groups: Delphi andprediction markets compared. *Foresight: The International Journal of Applied Forecasting* (8):17–20.

40. Wasserman, S, Faust, K (1994). *Social Network Analysis: Methods and applications.* (Cambridge University Press).

41. Barabási, A, Albert, R (1999). Emergence of scaling in random networks. *Science*. 286 (5439): 509–512.

42. Banerjee, A, Chandrasekhar, A G, Duflo, E, & Jackson, M O (2013). The diffusion of microfinance. *Science*, 341(6144): 1236498.

43. Holm, S (1979). A simple sequentially rejective multiple test procedure. *Scandinavian journal of statistics*, 6(2): 65-70.

44. Wilcox, R. R. (1996). Statistics for the social sciences. (Academic Press).

45. Cameron, A. C., & Miller, D. L. (2015). *A practitioner's guide to cluster-robust inference.* J Hum Resour, 50(2): 317-372.

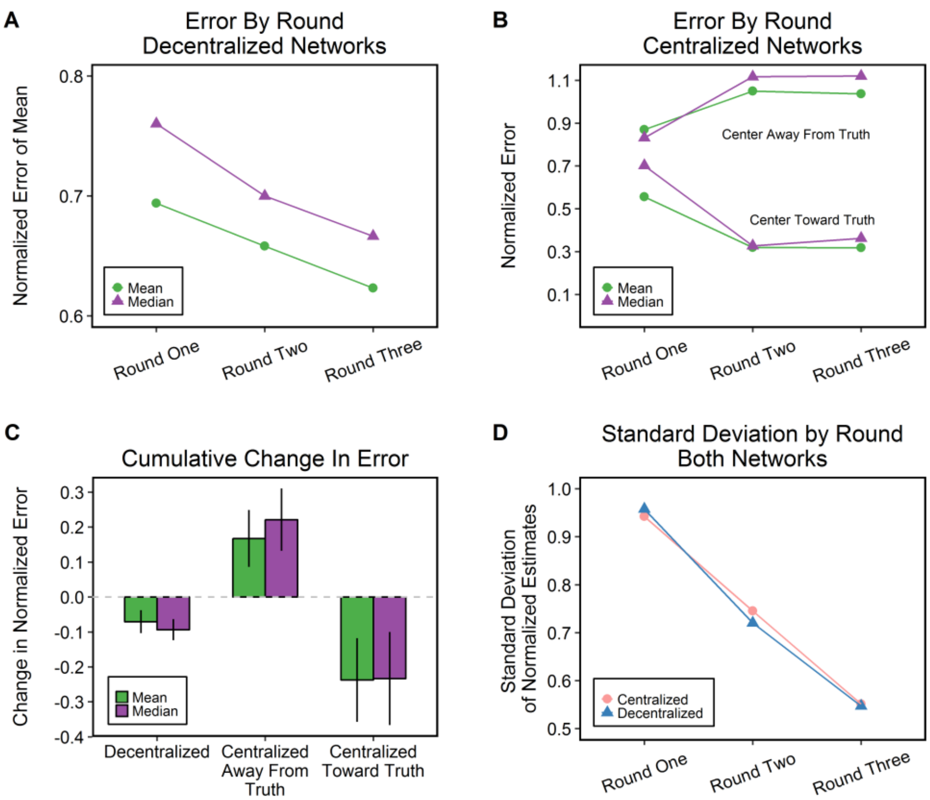

Fig. 1. Effect of social influence on group accuracy in centralized and decentralized networks. Average error and standard deviation in 13 experimental trials for each network condition. ***(A)*** In decentralized networks, both the mean and the median became more accurate over two rounds of social influence. ***(B)*** In centralized networks, the effect of social influence on the accuracy of the group mean and group median was determined by the initial estimate of the central node. Results are conditioned on whether the central node was in the direction of truth relative to the group estimate. ***(C)*** Total change from Round One to Round Three with bootstrapped 95% error bars, indicating that changes shown in panels A and B are significant. Both the mean and median of estimates in decentralized networks became more accurate (N=13, P<0.01 for mean, P<0.001 for median). For centralized networks, the mean and median became less accurate when the central node provided an estimate in the opposite direction of truth (N=13, P<0.01 for both mean and median). Both the mean and median became more accurate when the central node provided an estimate in the direction of truth (N=12, P<0.01 for the mean and median). ***(D)*** In both network conditions, the standard deviation in the distribution of estimates (i.e., diversity of opinions) decreased significantly after each round of revision (N=13, P<0.001 for both conditions).

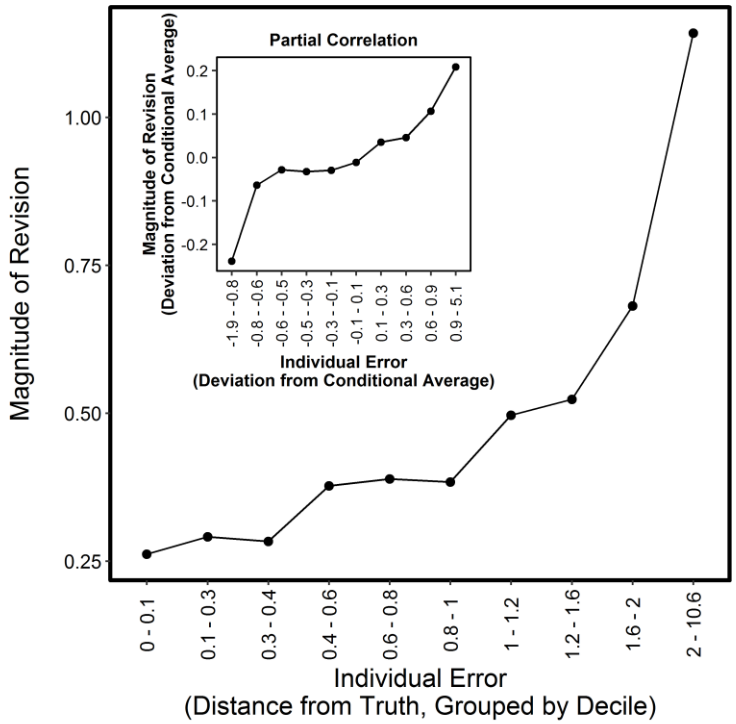

Fig 2. Correlation between revision magnitude and individual error. Each point in the main figure shows the average size of individuals' revisions from Round One to Round Three for individuals located in each decile of the distribution of individual error (i.e., average distance from zero error). Measured for N=4340 estimates provided by 1040 individuals assigned to one of 13 decentralized networks or 13 centralized networks. This figure shows a positive "revision coefficient," such that individuals with greater error in their initial estimates made significantly larger revisions. Controlling for correlation between estimates by the same individual (SI Appendix), we find a positive correlation between individual error and individual revision magnitude (N=4340, ρ=0.41, 95% CI [0.39, 0.43], P<0.001). **Inset:** On the y-axis, positive values indicate larger revisions than would be expected based on the distance between an individual's estimate and their neighborhood estimate. On the x-axis, positive values indicate greater initial error than would be expected given the distance between an individual's estimate and their neighborhood estimate. After controlling for the distance between each individual's initial estimate and the average estimate of their neighborhood, there is still a significant correlation between individual error and individual revision magnitude (N=4340, ρ=0.25, 95% CI [0.22, 0.28], P<0.001).

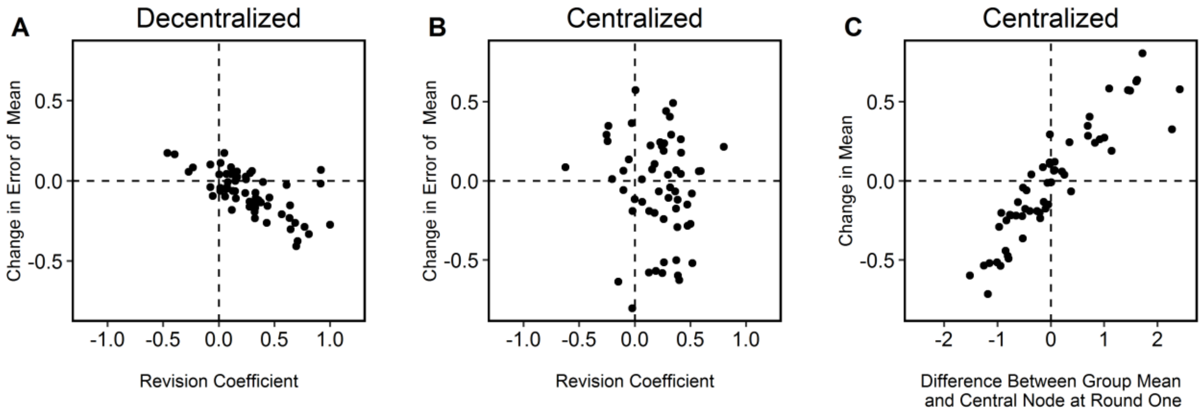

Fig. 3. Correlations with changes in group mean. Shown are all 59 estimation tasks completed over 13 experimental trials. In centralized networks, two estimation tasks are omitted where the central node did not provide any response. Decentralized networks show all 59 estimation tasks. ***(A)*** In decentralized networks, the "revision coefficient" for each group estimate – i.e., the partial correlation for all members of a network between individuals' accuracy and their revision magnitudes on a given estimation task – is highly correlated with the change in the error of the group mean (N=59, ρ=-0.71, 95% CI [-0.84, -0.51]). On estimation tasks in which groups exhibited larger revision coefficients, they showed significantly greater improvements in the accuracy of the group mean. **(B)** By contrast, in centralized networks, there was no significant correlation between the revision coefficient and the change in group mean (N=57, ρ=-0.16, 95% CI [-0.33, 0.10]). ***(C)*** In centralized networks, the change in the group mean is strongly correlated with the behavior of the central node. The difference between the initial group estimate and the initial estimate of the central node is highly correlated with the change in the group's estimate (N=57, ρ=0.92, 95% CI [0.88, 0.95]). When central node has an estimate larger than group mean, the group mean typically increased; when the central node is below the group mean, the group mean typically decreased.