Behavioral dynamics and influence in networked coloring and consensus

Stephen Judd, Michael Kearns1, and Yevgeniy Vorobeychik

Computer and Information Science, University of Pennsylvania, Levine Hall, Philadelphia, PA

Edited by Brian Skyrms, University of California, Irvine, CA, and approved July 16, 2010 (received for review February 3, 2010)

We report on human-subject experiments on the problems of coloring (a social differentiation task) and consensus (a social agreement task) in a networked setting. Both tasks can be viewed as coordination games, and despite their cognitive similarity, we find that within a parameterized family of social networks, network structure elicits opposing behavioral effects in the two problems, with increased long-distance connectivity making consensus easier for subjects and coloring harder. We investigate the influence that subjects have on their network neighbors and the collective outcome, and find that it varies considerably, beyond what can be explained by network position alone. We also find strong correlations between influence and other features of individual subject behavior. In contrast to much of the recent research in network science, which often emphasizes network topology out of the context of any specific problem and places primacy on network position, our findings highlight the potential importance of the details of tasks and individuals in social networks.

Related Work

Decentralized coordination is a problem of long standing interest. There are a number of game-theoretic models of coordination, in which players receive a positive payoff if and only if they can jointly coordinate on one of a collection of joint action choices. An alternative classic model is a cooperative game, which studies how players form coalitions. See Osborne and Rubenstein (4) for well known examples. Of central interest to us are the dynamics of the process by which players reach coordinated choice and the role of networks in this process; neither is informed by purely game-theoretic considerations.

While coordination games and cooperation in Prisoners Dilemma and other games have been extensively studied with human subjects over the years (5, 6–8), behavioral studies of coordination on networks are more recent. Kearns et al. (1, 9) study coloring and related problems on networks, although they do not focus on a particular parameterized family of networks as we do here. McCubbins et al. (10) and Kearns et al. (11) both observe that adding connections makes the coloring problem easier.

Experimental Methodology

The work described here continues a line of research at the University of Pennsylvania in controlled human-subject experiments on strategic behavior in social networks (1, 9, 13). In such experiments, each subject sits at a networked workstation with only a local (network neighborhood) view of the collective activity, and has a financial incentive to contribute to a global solution. In the coloring and consensus experiments we shall describe, the networks used in each experiment were chosen from a sto-

Author contributions: S.J. and M.K. designed research; S.J. and M.K. performed research; S.J., M.K., and Y.V. analyzed data; and S.J., M.K., and Y.V. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

1To whom correspondence should be addressed. E-mail: mkearns@cis.upenn.edu. Ordering of authors is alphabetical.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1001280107/-/DCSupplemental.
chastic generative model whose baseline network is a chain of six cliques (complete subnetworks) of six vertices each, for a total of 36 vertices. The model has a single parameter, a probability \( q \in [0,1] \). For any fixed value of \( q \), each edge not connecting two cliques in the baseline network is independently “rewired” with probability \( q \). To rewire an edge, one endpoint is randomly selected to be replaced by another, chosen uniformly at random from among the remaining 35 network vertices; thus regardless of the value of \( q \), the total number of edges is always preserved. See Fig. 1 for details and sample networks drawn from this stochastic model. Like many similar models from the literature (2), this model allows us to smoothly move between highly clustered, “tribal” network topologies with only loose interclique communication (\( q = 0 \)) to networks that are essentially random and unstructured, and dominated by “long-distance” links (\( q = 1 \)). Compared to earlier experiments (1, 9), which employed more heterogeneous network structures, our parametric model here also permits aggregating data from experiments with similar values of \( q \), which increases the statistical resolution of some of our tests.

For the experiments, we used values of \( q \in \{0,0.1,0.2,0.4,0.6,1\} \). These values were chosen so as to sample “interesting” networks with respect to a number of commonly studied structural properties of networks such as diameter and clustering coefficient; see Fig. 2. For each of these six values for \( q \), three independent networks were generated from the rewiring model, and each network used in one coloring and one consensus experiment; thus both tasks were performed on an identical set of 18 networks.

In the coloring experiments the collective objective was to color all vertices the same color within three minutes. In the coloring experiments, the collective objective was to color all vertices with a color that was different from those of any of its neighbors. Note that for coloring, there is a well defined minimum number of colors required for a solution to exist (the so-called chromatic number of the particular network in question), and this was the number provided in each experiment. For consensus, there is no obvious analogue, and one would generally expect the problem to become more difficult the more colors are provided; we arbitrarily allowed nine colors in each consensus experiment.

In the coloring experiments each subject received two dollars if a global (unanimous) consensus to any single color was reached, and zero dollars otherwise. In the coloring experiments subjects received two dollars if a valid global coloring was reached, and zero dollars otherwise. The subjects were given three minutes to reach a solution, although an experiment would terminate as soon as a solution was reached (thus, most experiments took less than three minutes). Each player was given only a local (neighborhood) view of the network. A screenshot of the Graphical User Interface is shown in Fig. 3.

All experiments were held in a single session lasting a total of several hours with 36 University of Pennsylvania students as subjects. Each experiment had a fixed network in which subjects were randomly assigned to vertices. The session was closely proctored and physical partitions were erected to ensure no communication between subjects, other than permitted by the system. In order to eliminate coordination based on previous experimental outcomes (particularly important for the consensus games), a random permutation between the colors chosen by a subject and the corresponding number of integer indices was chosen before each experiment, and the integer indices used to determine global state and solutions; thus what appears red to one player might appear blue to another, even though all players have a locally consistent view of the global state at all times. This randomization destroys the possibility of establishing social conventions across experiments, and renders any potential side communication about colors or buttons meaningless.

### Collective Behavior and Network Structure

Whether the two problems could be solved in the allotted time, and how long it would take to reach solutions, was a central question. All 18 coloring experiments ended with a global solution, as did 17 of the 18 consensus experiments. Thus overall collective performance was strong, a finding consistent with our previous experiments on other problems and network structures (1, 9).

Turning next to the effects of network structure, our first main result is shown in Fig. 4, which plots average solution time against the rewiring parameter \( q \). The key finding is that the coloring and consensus tasks induce qualitatively opposite collective performance as a function of \( q \): consensus appears to be much harder at low values of \( q \), and got easier as \( q \) increased. In contrast, the coloring problem was easiest at low \( q \). When the coloring and consensus solution times are aggregated separately for...
rather quickly converge on just two or three of the nine available colors, and most of the duration of the experiments involves a protracted network battle between the colors of different cliques, which each reach internal solutions rather quickly. A variety of interesting collective and individual behaviors can be seen in these visualizations; see the caption for detail.

**Individual Behavior and Influence**

Thus far we have reported performance results from a collective viewpoint, but ultimately collective behavior is a result of interactions between individual human subjects. Given the nature of our experiments, where a desired global state must be reached through local network interactions, it is natural to focus on measures of influence that subjects exert on their neighbors and on the collective outcome. The subject of individual influence is important enough to have spawned a substantial literature, particularly in the context of adoption cascades in networks (11, 12) where a decision maker is seeking out the most influential agents. Traditionally, the focus has been on identifying the relationships between influence and network position measures such as degree and centrality, viewing network position rather than individual human actors as the primary force. In contrast, below we argue that in our experiments it is *people* who are the primary conduits of influence in networks, and thus that human factors, often understudied in modern network science, may have much to do with influence and the success or failure of cascades.

We studied two distinct notions of influence in our experiments. The first, which we shall call *neighborhood influence*, attempts to measure the extent to which the actions of a player apparently cause changes in the actions of his neighbors. Fix a vertex $i$ in a given network, and suppose that at some time $t_0$ this vertex or player changes their color to $c$, and that their next color change to $c'$ comes at time $t_1 > t_0$. Then we will credit $i$ with an amount of influence equal to $1/(t_1 - t_0)$. The intuition here is that $j$ has adopted the current color of $i$, and that the amount of influence we credit $i$ with should diminish the longer this adoption took. Note that under this definition, $j$ could contribute more than one influencing event to $i$ (for instance, if $j$ toggled back and forth between $c$ and another color during the interval $[t_0, t_1]$), and the same color change by $j$ could contribute influence credit to more than one neighbor (for instance, if $j$ had another neighbor $i'$ who was also color $c$ at time $t$). However, neither of these phenomena diminishes the fact that the notion captures apparent temporal influences between neighboring players.

In order to filter out potentially coincidental color changes, we only credit $i$ in the manner described if $t_1 - t_0$ is at least $1$ s, an approximation of human reaction time combined with system display update latency (which was less than $100$ ms). Finally, we
define the overall neighborhood influence of a player in a given experiment to be the sum of all the influence credits as described, divided by its degree and number of color changes; thus the measure is already normalized to compensate for either a large number of neighbors or players who change color frequently.

Our second measure of influence is called outcome influence, and is simply the amount of time elapsed between a player’s final color change and the end of the experiment. We thus attribute the most credit to those players who anticipate the eventual global solution, and thus may have participated in its creation. Under this definition, the outcome influence of the last player to complete a global solution is zero. Note that neighborhood influence depends intimately on network structure, while outcome influence does not do so in any obvious or direct way.

Influence Variation by Subject and Position. Armed with these two notions of influence, our first interest is to examine whether the influence of an agent can be understood as arising primarily from their position in the network, or is to a significant extent a property of the actual human subject. In our experimental methodology, human subjects are randomly assigned to vertices at the beginning of each experiment. This policy helps to separate the effect of the human from the effect of the vertex (network position).

For each of our two notions of influence, we use a standard ANOVA test to detect whether variability of average influence among the 36 human subjects exceeds the variability of influence due to experimental conditions, such as random assignments of subjects to network vertices.

For neighborhood influence, we find that human subject variability is significant in coloring games (with $P < 0.001$), but not consensus games; while for outcome influence, human variation is significant with $P < 0.05$ in both game types. We conclude that overall, influence is not a product of just an agent’s network position, but arises from their intrinsic human behavioral patterns as well.

Correlates of Influence. Our two measures of influence quantify the (apparent) effects of a player or vertex on other players or vertices. We now examine a number of correlates of influence that are determined only by a player’s own actions, or their apparent reaction to their neighbors.

For instance, consider the change rate or instability of a player or position, which is the number of color changes per second in an experiment. For this measure we again find that human variation is significant under a standard ANOVA test, and furthermore that the change rate of a human subject is strongly negatively correlated with both their neighborhood and outcome influence in both game types (correlations ranging from $-0.40$ to $-0.51$ depending on the influence and game type, all with $P < 0.02$). We thus conclude that more stable players exhibit greater influence. Note that this finding is not obvious a priori, since neighborhood influence is already normalized for change rate, and for outcome influence the least stable players may still settle on their final color relatively early.

We also examined the notion of stubbornness, defined as the proportion of time that a player or vertex chooses a “noncompliant” color. In coloring, a color is noncompliant if there is another played by more neighbors. In coloring, a color is noncompliant if there is another played by fewer neighbors. As with change rate, we find stubbornness variation among subjects to be significant, and stubbornness is strongly negatively correlated with both influence types ($-0.45$ for neighborhood influence, $-0.50$ for outcome influence, both with $P < 0.02$) in consensus games (but not in coloring games). Thus it appears that rather than causing others to react to or coalesce around their noncompliant choices, stubborn players reduce their impact on their neighbors and the outcome of experiments. Despite the fact that human variation in influence cannot be explained by network position alone, and that there appear to be a number of other behavioral traits that strongly correlate with influence, it remains reasonable to ask what purely structural properties of network position still have some explanatory value.

The degree of a vertex is only weakly correlated with influence in consensus games ($0.21$ for neighborhood influence, $0.10$ for outcome influence, both with $P < 0.03$). Also, degree is not significantly correlated with either type of influence in coloring games. Vertex centrality (average inverse shortest path distance to all other vertices) is significantly correlated only with outcome influence in consensus games ($0.26$ with $P < 0.03$). The clustering coefficient of a vertex, which can be viewed as a measure of connectivity in a vertex’s neighborhood, is significantly correlated only with outcome influence ($0.086$, $P < 0.05$ in coloring, $-0.22$, $P < 0.001$ in consensus). Similarly, being one of the “connector” vertices between cliques in the baseline network before rewiring is correlated only with neighborhood influence in consensus games ($0.17$, $P < 0.03$). The evidence again suggests that influence seems more determined by the behavioral characteris-
tics of actual human subjects than by the abstract properties of network position.

An interesting related question is to what extent influence, as well as the measures of individual behavior above, correlate with the time the subjects took to reach a global solution. Stubbornness and change rate (averaged for each experiment) both exhibit strong negative correlation with running time in both games (correlations ranging from \(-0.81\) to \(-0.66\) with \(P < 0.005\)); stubbornness and instability seem to adversely affect decentralized coordination. On the other hand, we did not find significant correlation between running time and neighborhood influence in either game.

**Conclusion**

The results presented here have shown that simple and systematic changes in social network structure can elicit opposing effects on collective performance in two similar coordination tasks, and that the intrinsic behavioral traits of human actors may play a stronger role in their relative influence than purely structural properties of their network position. Both findings suggest the potential benefits of broadening the current emphasis of network science on topology. We would further note that our controlled experimental methodology—allowing us to systematically vary network structure across the same pool of human subjects—was crucial to the findings, since it placed the same subject in many different network positions throughout the session. Empirical field studies measuring only the structure and behavior within a single, fixed social network are handicapped in this regard.

**ACKNOWLEDGMENTS.** This research was supported by an Office of Naval Research MURI grant on network science.