



Organization Science

Publication details, including instructions for authors and subscription information:
<http://pubsonline.informs.org>

Facts and Figuring: An Experimental Investigation of Network Structure and Performance in Information and Solution Spaces

Jesse Shore, Ethan Bernstein, David Lazer

To cite this article:

Jesse Shore, Ethan Bernstein, David Lazer (2015) Facts and Figuring: An Experimental Investigation of Network Structure and Performance in Information and Solution Spaces. *Organization Science* 26(5):1432-1446. <https://doi.org/10.1287/orsc.2015.0980>

This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. You are free to download this work and share with others for any purpose, even commercially if you distribute your contributions under the same license as the original, and you must attribute this work as “*Organization Science*.” Copyright 2015 INFORMS. <http://dx.doi.org/10.1287/orsc.2015.0980>, used under a Creative Commons Attribution License: <http://creativecommons.org/licenses/by-sa/4.0/>.

Copyright © 2015, INFORMS

Please scroll down for article—it is on subsequent pages



With 12,500 members from nearly 90 countries, INFORMS is the largest international association of operations research (O.R.) and analytics professionals and students. INFORMS provides unique networking and learning opportunities for individual professionals, and organizations of all types and sizes, to better understand and use O.R. and analytics tools and methods to transform strategic visions and achieve better outcomes. For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Facts and Figuring: An Experimental Investigation of Network Structure and Performance in Information and Solution Spaces

Jesse Shore

Information Systems Department, Questrom School of Business, Boston University, Boston, Massachusetts 02215, jccs@bu.edu

Ethan Bernstein

Organizational Behavior Unit, Harvard Business School, Boston, Massachusetts 02163, e@hbs.edu

David Lazer

Department of Political Science, Northeastern University, Boston, Massachusetts 02115, d.lazer@neu.edu

Using data from a novel laboratory experiment on complex problem solving in which we varied the structure of 16-person networks, we investigate how an organization's network structure shapes the performance of problem-solving tasks. Problem solving, we argue, involves both *exploration for information* and *exploration for solutions*. Our results show that network clustering has opposite effects for these two important and complementary forms of exploration. Dense clustering encourages members of a network to generate more diverse information but discourages them from generating diverse theories; that is, clustering promotes exploration in information space but decreases exploration in solution space. Previous research, generally focusing on only one of those two spaces at a time, has produced an inconsistent understanding of the value of network clustering. By adopting an experimental platform on which information was measured separately from solutions, we bring disparate results under a single theoretical roof and clarify the effects of network clustering on problem-solving behavior and performance. The finding both provides a sharper tool for structuring organizations for knowledge work and reveals challenges inherent in manipulating network structure to enhance performance, as the communication structure that helps one determinant of successful problem solving may harm the other.

Keywords: networks; experiments; clustering; problem solving; exploration and exploitation; knowledge; information; communication; search

History: Published online in *Articles in Advance* April 27, 2015.

Introduction

How does the clustering of organizational and social networks affect problem-solving behavior and performance? Unfortunately, answers to that question remain incomplete. Substantial recent research implies that clustering—the degree to which people with whom a person is connected are themselves connected to each other—can improve problem-solving performance by increasing coordination (e.g., [Kearns et al. 2006](#), [McCubbins et al. 2009](#)), supporting the managerial trend toward increasing connectedness in workplaces worldwide ([Bernstein 2012](#)). By contrast, equally powerful research suggests that clustering can undermine performance by fostering an unproductive imbalance between exploration and exploitation, even for simple tasks ([Lazer and Friedman 2007](#), [Mason et al. 2008](#), [Mason and Watts 2012](#)). A generalized net effect of clustering on problem-solving performance remains unresolved.

We move a step closer to resolving the question above by presenting new evidence from a laboratory experiment that unites these disparate findings under a single theoretical roof. Borrowing terminology from [March \(1991\)](#), we

hypothesize that effective problem solving requires both *exploration of information space* (for facts that may be important pieces of the puzzle) and *exploration of solution space* (for theories, or interpretations of facts, that combine puzzle pieces into an answer). Our reading of the literature suggests to us that the types of communication network structure that support exploration of information space may not be those that support exploration of solution space. We therefore adopt a novel, data-rich experimental platform that emphasizes verisimilitude: subjects complete a collective problem-solving task that people might confront in real organizational settings, which requires both exploration of information space and exploration of solution space to solve.

We find that clustering promotes exploration through information space but inhibits exploration through solution space. Through the active communication of information, individuals in a connected cluster tend to be in possession of the same knowledge and to be aware of each other's theories (solutions). The mutual knowledge facilitates an efficient search for additional information, but the mutual awareness of each other's theories results

in a convergence in interpreting that information, reducing the exploration of theory space.

The same network structure, therefore, can either promote or inhibit knowledge diversity, depending on whether that knowledge consists of *information* or *interpretations of information*. The implication is that “good” communication structures may only be good for parts of the collective problem-solving process: a structure that improves performance now may degrade performance later.

Differentiating Information and Solution Spaces

Overview: Clustering, Mutual Knowledge, and Performance

Individuals in densely clustered networks accrue shared mutual knowledge by communicating with each other (e.g., [Granovetter 1973](#), [Hansen 1999](#), [Burt 2004](#)). How clustering affects problem-solving performance, however, remains an open question. Although this is a hot topic that has attracted high-quality scholarship, inconsistent performance results make conclusions elusive.

When such inconsistency persists, a finer categorization of the observed phenomenon can help to resolve contradictions and open a path for progress ([Christensen and Carlile 2009](#)). In a detailed reading of prior work on the relationship between clustering, mutual knowledge, and problem-solving performance, we surface patterns indicating that clustering may have different consequences for exploration for information (facts) and exploration for solutions (interpretations of facts, or theories). We organize the following review of the literature accordingly, concluding with testable propositions.

Exploring Solution Spaces: Clustering Undermines Performance?

Since [March \(1991\)](#), problem solving has been seen to require some degree of “exploration” of an unknown landscape of possible solutions. [March \(1991, pp. 71 and 85\)](#) contrasts exploitation (“such things as refinement, choice, production, efficiency, selection, implementation, execution” from which “returns are positive, proximate, and predictable”) with exploration (“captured by terms such as search, variation, risk taking, experimentation, play, flexibility, discovery, innovation” from which “returns are uncertain, distant, and often negative”). How does clustering affect the balance between exploration and exploitation? Clustering has been connected with reduced exploration and thus depressed performance in problem-solving networks.

To study the connection between clustering and exploration of solution space, building on [March’s \(1991\)](#) dichotomy, network science researchers (e.g., [Lazer and Friedman 2007](#), [Mason et al. 2008](#), [Mason and Watts 2012](#)) have adopted the pragmatic view that exploration

is developing new solutions that previously did not exist in the organization, whereas exploitation is converging on (adopting/copying) preexisting solutions already at hand within the organization. Implementing the best solution to a problem involves both, and therefore structures that enable optimal problem solving need to effectively balance biases toward one or the other.

Unproductive biases in collective problem solving favoring too much exploitation or exploration can take familiar forms. For example, an individual who copies a neighbor’s answer, solution, or theory is probably doing so because she or he expects the returns to be more “positive, proximate, and predictable” than the more “uncertain, distant, and often negative” returns of trying to solve the problem alone, or because the neighbor’s choice of that answer seemingly provides social proof of its value ([March 1991, p 85](#)). Even if exploitation includes not just copying exactly but also refinement of the copied solution (e.g., [Anjos and Reagans 2013](#)), such exploitation can result in premature convergence (a reduction of solution diversity) at the collective level, thereby reducing performance.

For example, [Lazer and Friedman \(2007\)](#) use agent-based modeling to show that agents in highly connected networks (i.e., networks in which the average length of the path between individuals is short) converge rapidly on a relatively good solution as they adopt the solutions of their neighbors. Highly connected networks (like a complete cluster) are thus “efficient” in their ability to facilitate diffusion of good solutions among network members. In the short run, they outperform less connected networks, whose members are less likely to be aware of good solutions elsewhere in the network. It is in the less connected—and thus less efficient—networks, however, that we will find individuals who are not yet exploiting the current best solution conducting more exploration of the solution landscape and bringing more potential solutions into the network. Indeed, agents in [Lazer and Friedman’s](#) inefficient networks eventually converged on better solutions, collectively, than agents in efficient networks.

Substituting simulated agents for real human subjects, a recent large experiment ([Mason and Watts 2012](#)) replicated the finding that networks that collectively explored more did better in the long run. But [Mason and Watts’s \(2012\)](#) more efficient networks explored more, appearing to contradict [Lazer and Friedman’s \(2007\)](#) finding that less efficient networks explored more. The two studies are more consistent, however, if they are interpreted in terms of *clustering* instead of efficiency. In both studies, clustering is associated with less exploration.

How does clustering suppress exploration of solution space? The effect of local clustering in suppressing exploration in [Mason and Watts \(2012\)](#) may be due to a process of complex contagion ([Centola and Macy 2007](#), [Centola 2010](#)), in which adopting a neighbor’s solution is

more likely to occur within, rather than between, clusters of ties. Novel exploratory solutions both are uncertain in advance and have material consequences—good reasons to adopt a solution on which other people seem to have already reached a consensus. In the language of March (1991), this is to say that exploration of solution space should be less extensive within clusters; Mason and Watts show strong evidence corroborating that insight. Thus, studies focusing on how clustering affects the exploration of solution space have found a dampening effect: clustering unproductively biases individuals away from the exploration of new solutions and toward the exploitation of existing ones.

Exploring Information Spaces: Clustering Aids Performance?

In contrast to studies of solution spaces, studies of how individuals explore *information* spaces to find the facts needed to solve problems suggest that clustering may increase exploration and therefore increase performance. When the target of exploration is new information (facts) rather than new solutions (figuring), clustering appears to have the opposite effect.

For example, there is an expanding literature on collective problem solving in networks that uses variants of the “distributed graph-coloring problem” as an experimental task (e.g., Kearns et al. 2006, Kuhn and Wattenhofer 2006, Kearns et al. 2009, Judd et al. 2010). In graph-coloring tasks, subjects must choose from a discrete set of colors such that they do or do not (depending on the task) match the choice of their neighbors. No subjective interpretation is required: each subject takes in information about his or her neighbors and selects a color according to the instructions. The lack of interpretation makes these tasks quite different from those described in the previous section. The critical variable for success in these tasks is not experimentation but rather coordination.

For our purposes, it is unfortunate that these high-quality graph-coloring studies largely investigate the effects of density rather than clustering. Although their results cannot be assumed to apply exactly to clustering, density and clustering are related measures: clustering is the degree to which all the individuals within a local neighborhood are connected with each other, and density is the analogous property of the whole network.

In general, greater density of ties improves performance in graph-coloring tasks (McCubbins et al. 2009). A major mechanism appears to be greater mutual knowledge; knowing what one’s neighbors’ neighbors are doing makes the distributed graph-coloring problem much easier to solve (Enemark et al. 2014). Clustering of ties means that many of one’s neighbors’ neighbors are also one’s own neighbors, and therefore that there is extensive mutual information in one’s neighborhood that can increase performance in exploration through information space.

That finding, however, appears to contradict some of the earliest laboratory-based social network experiments, conducted by Bavelas (1950) and Leavitt (1951) at the Massachusetts Institute of Technology’s Small Group Network Laboratory on five-person networks. In these experiments, five-person star graphs (centralized unclustered networks) were more effective in collective problem solving on information-focused tasks than five-person complete cliques (maximally clustered networks) (Bavelas 1950, Leavitt 1951), because the star graph all but guarantees that at least one “central” person receives all the information available in the network, which he or she can then disseminate to the other members of the organization (Guetzkow and Simon 1955). Although there was some evidence that the effect of clustering was contingent on the difficulty of the problem-solving task (Shaw 1954), it was later established that, ultimately, centralized (unclustered) networks performed better for both simple and difficult problems once a centralized and coordinated decision structure evolved within the experimentally imposed communication structure (Mulder 1960).

In sum, recent experiments show better-coordinated exploration of information space in clusters, whereas the early experiments show the opposite. One way to resolve this contradiction is by looking more closely at the characteristics of the information-focused task. Whereas multiple network ties could be used simultaneously in more recent experiments, only one tie could be used at a time in the early experiments. In other words, in the earlier experiments, the connection between clustering and shared mutual knowledge was broken. Indeed, the more paths information could take through the network, the less certain participants could be that they were communicating in an efficient way to complete the task. Although one can find real-world collective problem-solving tasks that are similar to the Bavelas (1950) experiments, these seem to be much more the exception than the rule. As long as communications can be addressed to more than one person at a time—for example, at a meeting or on a social media platform—then clustering would ensure that those individuals had full access to shared information, which is argued in both recent and early experimental work to be associated with better performance.

Although these earlier results appear to diverge from those of the recent graph-coloring experiments, the differences are probably due to differences in the experimental protocols described above. Taking the protocols into account, both early and recent experiments show that network structures that promote full access to information also enable more coordinated problem solving.

Summary: Clustering and Problem-Solving Performance in Solution and Information Spaces

Taken together, these streams of research on exploration for solutions and exploration for information provide

Figure 1 Impact of Clustering on Performance by Domain

DEGREE OF CLUSTERING	Dense	Exploration (coordinated)	Copying, imitation
	Sparse	Duplication, redundant effort	Exploration (uncoordinated)
		Information space	Solution space
		DOMAIN	

substantial evidence that the mutual knowledge enabled by clustering simultaneously enables *coordination* of fact finding (in information space) and *copying* of interpretations (in solution space), as summarized in Figure 1.

When everybody knows what everybody else is doing, it allows each individual to act in a way that is complementary to the actions of the rest of the group; a lack of clustering would therefore be more likely to produce noncomplementary duplicated work. Similarly, the greater the degree to which individuals are aware of each other's interpretations, as would be the case within a cluster (Centola and Macy 2007, Aral and Van Alstyne 2011), the more likely they are to copy a consensus view rather than fully exploring the entire space of possible solutions to derive their own; a lack of clustering would instead be associated with more independent, uncoordinated interpretation of the facts. Although clustering does not guarantee either outcome, it does—by promoting mutual awareness—make it easier for individuals to both coordinate and copy in order to avoid redundancy.

An alternative way to frame the existing literature, as suggested by Figure 1, is that different types of exploration are valuable in different domains. In the exploration of information space, coordinated exploration may be valuable because it avoids duplication of work. In the exploration of solution space, uncoordinated exploration may be valuable because it avoids copying, which could lead to premature convergence on a suboptimal solution. Unfortunately, clustering simultaneously affects coordination and mutual knowledge in both domains.

To illustrate the difference between the effects of clustering in the two domains, consider the case of market analysts making forecasts. First, they need to gather a range of data: the unemployment rate, gross domestic product growth, the availability of credit, costs of inputs to production, etc. If analysts have a body of mutual and shared information, they will be less likely to collect data that are already known and will instead gather new information to improve their forecasts. Mutual and shared knowledge promotes greater exploration of information

space as a consequence of emergent coordination. However, it takes a great deal of interpretation to turn this information into a forecast. The more mutual and shared knowledge the analysts have, the more likely they are to be influenced by each other's interpretations of what the data mean. For example, prior to the financial crisis of 2008, there was a shared interpretation among analysts that the amount of leverage in the U.S. mortgage market was sustainable (Lewis 2010). Mutual and shared knowledge inhibits exploration of solution space (in this scenario, analysts can be said to be “solving” the puzzle of what the future state of a market will be).

In sum, our reading of the literature on clustering suggests the following possible interpretation: the most extensive aggregate exploration in solution space occurs when actors are independent (uncoordinated), whereas the most extensive aggregate exploration in information space occurs when actors are interdependent (coordinated). Our novel experiment, by adopting a more complex task and paradigm, allows us to evaluate whether clustering affects performance in information space and solution space differentially. As shown in Figure 1, we expect to see clustering associated with more exploration in information space but with less exploration in solution space.

Data and Methods

The Experimental Platform

The Task. To instrument the connection between clustering, solution exploration, information exploration, and problem-solving performance, we aimed to develop an experimental platform with several key characteristics: (a) maximum verisimilitude, which means both that the task was similar to real problem-solving work and that the means for accomplishing the task within the platform had real-world analogues; (b) maximum accessibility, which required the task to be easily understandable and solvable with expertise commonly available in our subject pool; and (c) maximum instrumentation, which required that actions taken by the participants be captured as richly as possible in subsequently analyzable data.

Based on these criteria, we selected a whodunit protocol, much like a game of Clue® or Cluedo®, in which the task involved piecing together clues, or facts, to “connect the dots.” Our task therefore bore some resemblance to the common murder mystery protocol in group research (Stasser and Stewart 1992, Stasser and Titus 2003), but with the following key differences in order to instrument our research question: (a) as clustering of communication ties would be a variable of interest (see the Treatments section below), our organizations would not be fully interconnected, as opposed to groups, which have a density of 1.0 by definition; (b) to accomplish such variance in clustering, our organizations would consist of 16 members rather than groups

of 3 or 6; (c) to model exploration (rather than only discussion), some clues would not be distributed to members of the organization but instead be accessible via a Google-style search so that access to new information was limited not by what other members knew but by the questions an individual member asked of the search engine; and (d) because search and sharing were enabled by technology (e.g., Dennis 1996), exploration for information was less limited by synchronous airtime in discussion (that is, just because one person was sharing did not mean that others could not do so simultaneously, as would be the case in a meeting). These differences do not reflect a critique of the original murder mystery protocol but rather reflect our interest in a substantially different question about clustering in networked knowledge-centered organizations.

Rather than creating a platform entirely from scratch, we were invited to customize a platform developed by the U.S. Department of Defense's Command and Control Research Program called ELICIT (Experimental Laboratory for Investigating Collaboration, Information-sharing, and Trust), which already had many of the characteristics we sought. Although we modified much of the platform, we agreed to preserve the nature of the Department of Defense's whodunit task, which involved predicting the who, what, where, and when of an impending terrorist attack (in place of, for example, the who, what, and where of a murder in Clue). More detail is available in Online Appendix A (available as supplemental material at <http://dx.doi.org/10.1287/orsc.2015.0980>).

Specifically, participants were faced with four logically independent subproblems to solve: (a) who would carry out the attack (group involved), (b) what would be the target (e.g., an embassy or a church), (c) where the attack would take place (country), and (d) when the attack would take place (with four interdependent components—month, day, hour, and time of day (A.M. versus P.M.)). Each question and subquestion had a dedicated text box in which to register an answer. Participants had 25 minutes to solve the problem.

Subjects could choose from a discrete set of actions. They could search for new clues by entering a keyword into a search text box and clicking a button; the tool would search for that keyword from among the 69–83 facts contained in the predetermined factoid set for that problem. They could share clues they already had (one at a time) with one or more of their neighbors (other subjects with whom they shared a network tie) and, if they wished, add free-text annotations to these shared clues. They could register their theories by typing them into the separate spaces given for the who, what, where, and when subproblems. Finally, they could check their neighbors' registered solutions at any time and any number of times.

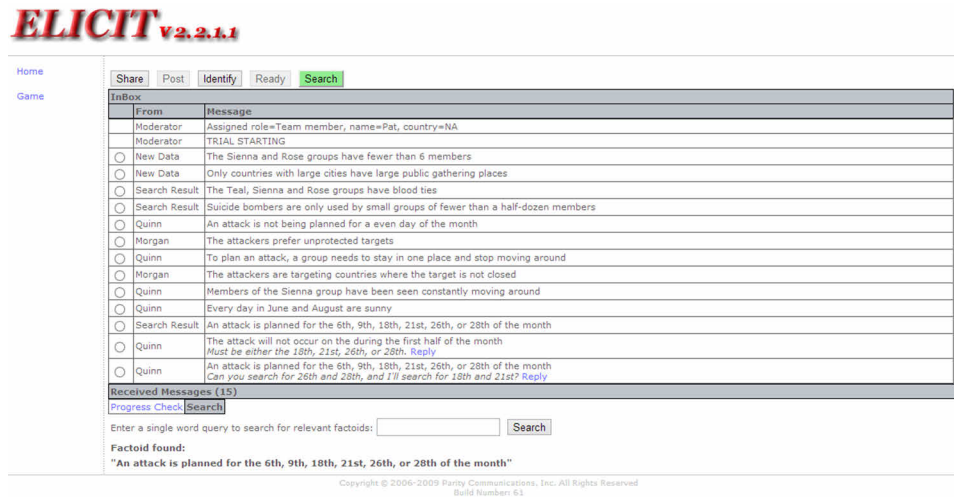
Participants were rewarded 15¢ per minute per subproblem (3.75¢ per minute for each component of

“when”) for which they had a correct answer registered, for a maximum of 60¢ for each minute (equivalent to \$36 dollars per hour). A participant therefore had a strong incentive to record a theory as soon as he or she had developed it and to adjust it as soon as new information warranted. Although success required help from neighbors—through sharing clues, annotating clues, and viewing neighbors' registered solutions—we chose to provide incentives exclusively at the individual level. Competition or group-level rewards would have introduced new interdependencies into the data that would have complicated interpretation of the interdependencies of primary interest: those created by network structure. The interaction of network structure and such interdependencies in incentives is indeed an important topic of study, but, for this experiment, we limit ourselves to the primary phenomenon. Since the information space was large and time was limited to 25 minutes, it was virtually impossible for people to find all the clues necessary to solve the problem alone, ensuring that they would share information (as they did) without the need for group-level incentives. At no point during the experiment did anyone know for certain whether his or her answers were correct, just as would be the case in real life.

Execution of the Experiment. The experiment was carried out in a laboratory setting, with each participant seated at a computer in a private carrel. All experimental activities were executed through a Web browser interface (see Figure 2), with the exception of scratch paper, which was collected and scanned at the end. Each experimental run lasted 25 minutes. Participants were given two clues at the start of each round and were allowed to search for more clues once per minute. The initial distribution of clues was not correlated with outcomes. Each clue was only relevant to one subproblem. Some clues contained useless or misleading information. Subjects had to combine multiple clues to conclusively arrive at the correct answer. The number of clues necessary to solve a single subproblem ranged from 2 to 10, with a median of 5 and a mean of 5.3.

Each experimental session began with an instructional video explaining the platform and the task in uniform fashion to every subject across all sessions. To control for individual aptitude, each person then took a pretest with the same format as the experimental task, with a similar but smaller-scale problem and without interaction with other participants.

After the pretest, subjects could take part in up to three runs of the experiment. Within any given run, all subjects had the same level of experience with the experiment. That is, if a subject were taking part in a second run, then all of the other subjects in that run would also be taking part for the second time. Holding this experience constant eliminated the possibility of spurious correlations between learning and network structure. In the

Figure 2 (Color online) Screenshot of Web Browser Interface Used in the Experiment

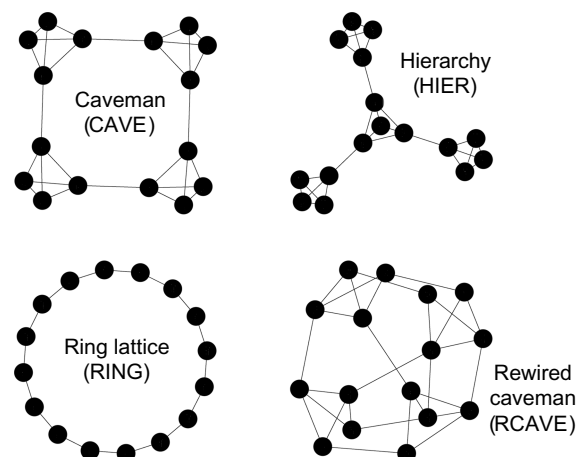
multivariate analyses, we include a round variable that indicates the experience of the participants in number of runs through the experiment.

At any given time, lab space allowed us to run up to two concurrent 16-person experimental runs. Three different problem sets were used to limit contamination between sessions of the experiment. In the results, these are indicated by the factoid set variable. Additionally, within each problem set, proper nouns (names of places, countries, and terrorist groups) were randomly permuted to reduce the risk of contamination between sessions.

Before each run began, a network treatment was chosen at random and study subjects were randomly assigned to a position in that network that was uncorrelated to their physical location in the laboratory. Participants were assigned a pseudonym to further obscure their identities from each other; the pseudonyms were shuffled before each round.

There were 417 unique individuals; they played a total of 1,120 person-rounds. Participants were recruited through the subject pool of a large university in the northeast United States. The mean self-reported math and verbal SAT scores were 716 and 701, respectively, consistent with reported data from the university. The self-reported genders were 49.5% male and 50.5% female. This paper reports results from a subset of the collected data, consisting of 816 person-rounds played by 352 unique individuals. The remaining data included additional treatment variables intended to test other phenomena and are not comparable to the data we analyze here.

Treatments. We tested four 16-person network treatments (see Figure 3 for visualizations and Table 1 for descriptive statistics), within which subjects were randomly assigned to network positions. At the top left of Figure 3 is the “caveman” (CAVE) network (Watts 1999), containing four four-person cliques. The “hierarchy”

Figure 3 Visualizations of the Network Treatments

(HIER) network is likewise composed of four such cliques, but it is arranged in a conventional centralized structure. The “rewired caveman” (RCAVE) is a small-world network, constructed by removing links from the caveman network, then adding links that create shortcuts through the network. Members of the rewired caveman network are therefore “closer together” topologically: the most distant pair of individuals is only three hops

Table 1 Descriptive Statistics for Networks

	CAVE	RCAVE	HIER	RING
Average degree	3.5	3.625	3.375	2
Density	0.467	0.483	0.45	0.267
Average path length	2.47	1.99	2.81	4
Diameter	5	3	5	8
Mean clustering coefficient	0.667	0.304	0.727	0
Centralization (eigenvector)	0.033	0.115	0.161	0

away, and the average distance between all pairs is shorter than it is in the caveman and hierarchy networks. The rewired caveman network is also more centralized and less clustered than the caveman network. Finally, there is the “ring” (RING) network, which is neither clustered nor centralized.

The network ties imposed through the treatments allow communication of clues and theories. With respect to clues, the average subject used all available network ties 93% of the time when they shared information and added annotations to 63% of those shared clues, suggesting that the opportunities for communication were realized almost fully in practice and that the essential structural nature of the network treatments was preserved.

In some of our statistical analyses, rather than testing the effects of the network treatment as a whole, we tested the effects of nodal degree and clustering coefficient. Both are individual-level structural metrics: degree is simply the number of connections a node has; the clustering coefficient—the number of existing connections among neighbors divided by the number of possible connections among neighbors—measures the extent to which a node’s neighbors are also neighbors of each other (Watts and Strogatz 1998).

Outcome Variables

We treated exploration and exploitation as mutually exclusive classifications of a single action, as has been customary (March 1991, Lazer and Friedman 2007, Mason and Watts 2012). That is, for the purposes of this experiment, if an action is an example of exploration, then it is not also an example of exploitation. This was important in our measurement of these constructs, because in different cases, one or the other would be easier to observe. For example, it would be easier to measure the amount or extent of information gathering (exploration) than the amount or extent of mental processing of information already held (exploitation). Additionally, because we were working at both the collective and individual levels, we specified measurements of exploration and exploitation at each level separately.

Exploration in Information Space. In our experiment, information space was explored by searching for facts. We measured exploration in information space at the collective level with two variables: the number of unique facts discovered by the group as a whole and the ratio of total facts found to unique facts found. The latter measure can be interpreted as the degree to which facts were found multiple times within the same group, or simply as the redundancy of facts found. The more unique facts and the lower the redundancy, the greater the exploration of information space.

At the individual level, we measured the total number of facts found by (a) the subject’s own search (the search interface did not return facts that an individual already

possessed, so the total number of facts represented the extent of exploration in information space by search) and (b) the redundancy of facts received from neighbors (the ratio of total facts received to unique facts received).

Exploration and Exploitation in Solution Space. We measured exploration in solution space at the collective level in terms of the total number of unique theories that were registered during the experiment. At the individual level, we focused on exploitation because it is easier to observe. Specifically, exploitation in solution space took the form of checking and then copying a neighbor’s theory. We had time-stamped records of every action undertaken during the experiment, so checking neighbors’ theories could be directly measured from the data. We define *copying* as an individual checking a neighbors’ answers and then registering one of those answers the next time that individual enters his or her own theories, provided this occurs within 10 minutes of the original observation of the neighbor’s answer.

Establishing the uniqueness of theories required us to consolidate answers such as “power plant,” “power-plant,” and “electric power plant” into one theory, which we did in two steps. First, automated preprocessing of entries removed punctuation, converted the text to lowercase, and combined repeated entries where one example had a simple typo (defined as a single insertion, substitution, or deletion of a character other than the first letter of the word). Second, we used a human coder to remove more substantial typos (such as transposing letters or whole phonemes) and to combine answers in which the intent was clearly the same (e.g., we considered “power plant” and “electric power plant” to be the same).

Performance. Given that clustering was expected to have both positive and negative effects, depending on the domain of reference, we also measured overall performance. Performance was measured in pay per minute received by individuals. When measuring the performance of an entire network, we simply added up all the members’ pay per minute.

Statistical Framework

Wherever possible, we considered both individual-level and collective-level correlations. At the collective level, we had 51 independent data points, each corresponding to a run of the experiment. For these models, we used ordinary least squares (OLS). At the individual level, we had 816 observations with two types of interdependence. First, they were nested into the 51 runs, mentioned above. Second, since individuals played multiple runs of the experiment, the 816 observations were generated by 352 unique individuals. We therefore included random effects for both run and unique individual and estimated linear mixed models (LMM).

For discrete outcome variables at the individual level, we used mixed-effects Poisson (GLMM-Poisson)

regression (Bates et al. 2012). Because the number of times a participant checked neighbors' answers exhibited zero inflation, we estimated a mixed-effects zero-inflated Poisson model (GLMM-ZIP) in a Markov chain Monte Carlo framework (Hadfield 2010). Statistical analysis was carried out in R (R Core Team 2012). Analytical code and data are available at Dataverse (<http://thedata.org>).

The tables of results that we present in the following section contain evidence to support causal and noncausal inferences. The experimental treatments we imposed are the relational structures into which individuals are placed: the networks as a whole. Aggregate results at the collective level can therefore be inferred to be causal results of these treatments. For example, if performance in the ring network were lower than performance in the hierarchy, we could conclude that this was due to the network structure as a whole. Results from statistical models at a more microscopic level that use individual-level variables—even exogenously imposed structural variables, such as degree and clustering—should be interpreted as correlations rather than as evidence of causality. Outcomes at the individual level depend not only on the local structure but also on the structure of the remainder of the network. Although node-level metrics do not account for this, they can help us understand the causal results at the collective level by providing more insight into individual behavior.

The next section reports on each of the three sets of variables: exploration of information space, exploration of solution space, and performance.

Results

Our results provide evidence for the proposition we outline above: clustering promotes exploration of information space and inhibits exploration of solution space. After presenting that evidence, we report additional findings that the clustered networks outperformed the unclustered networks in this particular problem-solving task.

Information Space

We find that clustering was associated with greater exploration of information space in the sense of finding and sharing not more but rather more unique (or less redundant) information. By being in a cluster, individuals tended to contribute more to the *collective* exploration through information space—not from more search, but rather by being more coordinated in their search. We find evidence of this at both the individual and collective levels.

Clustering does not lead to a larger number of searches, as evidenced by two findings. At the individual level, the clustering coefficient of an individual's position was not correlated with the number of searches for facts performed (see Table 2, Model 1). Likewise, at the collective level, the clustered networks did not search at a different rate than the rewired caveman did (see

Table 2 Exploration and Exploitation in Information Space

DV	Total facts found by search				Total searches ÷ Unique facts found		Total receipts ÷ Unique facts received	
	1	2			3		4	
Model number								
Random effects	Variance				Variance			
Individual	0.012						0.001	
Run	0.000						0.006	
Fixed effects	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>
Intercept	3.085	< 0.001***	267.037	< 0.001***	4.002	< 0.001***	0.512	< 0.001***
Degree	−0.085	< 0.001***					0.161	< 0.001***
Clustering coefficient	−0.019	0.423					−0.131	< 0.001***
HIER			2.424	0.722	0.114	0.198		
RCAVE			7.231	0.257	0.199	0.017*		
RING			46.248	< 0.001***	0.761	< 0.001***		
Pretest	0.002	0.025*					−0.000	0.315
Second round	0.061	0.002**	18.668	0.003**	0.137	0.082	0.064	0.031*
Third round	0.071	< 0.001***	21.320	< 0.001***	0.190	0.018*	0.112	< 0.001***
Factoid set 2	0.119	< 0.001***	35.172	< 0.001***	−0.236	0.004**	−0.116	< 0.001***
Factoid set 3	0.098	< 0.001***	28.814	< 0.001***	0.075	0.353	−0.047	0.117
<i>n</i>	816		51		51		816	
Model type	GLMM (Poisson)		OLS		OLS		LMM	
Unit of analysis	Individual		Aggregate		Aggregate		Individual	

Note. DV, dependent variable.

p* < 0.05; *p* < 0.01; ****p* < 0.001.

Model 2; note that the caveman network serves as the reference category).

Although clustering does not increase the number of searches, we found that it nonetheless increases exploration of information space by promoting less redundant search. At the collective level, this is demonstrated by the finding that the facts found by the caveman network were significantly less redundant than those found by the ring and rewired caveman networks (Model 3). In other words, the caveman network was more efficient in exploring information space, in that subjects collectively covered more new ground with each search. The mean redundancy of facts found by the hierarchy network was also lower than that of the ring and rewired caveman networks, although the difference was not significant. Exploration of information space at the collective level requires not only that information be found but also that it be effectively transferred throughout the network; we found that here, too, clustering increased exploration. Individuals in clustered positions received significantly less redundant information from their network neighbors (Model 4).

Degree also had an effect on exploration of information space, as it was associated with fewer searches (Model 1). At the collective level, the ring network searched at a higher rate than the other three networks (Model 2), which was likely a consequence of its members having a low degree.

Solution Space

In solution space, we find that clustering inhibits exploration by promoting the copying of neighbors' answers (see Table 3). We measure these effects in terms of the propensity of individuals to check and copy their neighbors' theories and on the aggregate amount of copying and the total number of unique theories registered at the whole-organization level. Both the checking and copying of neighbors' theories indicate less extensive exploration of solution space, and we find that clustering at the individual level is associated with more checks of neighbors' theories (a marginally significant finding; see Model 5) and more outright copying of their theories (Models 6 and 8). Moreover, conditional on copying a neighbor's theories at all, those in clustered positions were more likely to copy an *incorrect* theory (Model 7), whether the person doing the copying originally had the correct answer, an incorrect answer, or no answer at all. At the collective level, the two clustered networks had significantly fewer unique theories registered in the aggregate than did the unclustered networks (Model 9). In other words, it appears that whereas clustering increased exploration of information space, it inhibited exploration of solution space.

Additionally, consistent with the predictions of the information-processing literature that information overload

Table 3 Exploration and Exploitation in Solution Space

DV	Theory checking		Theory copying		Incorrect copying		Theory copying		Unique theories	
Model number	5		6		7		8		9	
Random effects	Variance		Variance		Variance					
Individual	0.102		0.028		0.200					
Run	0.000		0.035		0.039					
Fixed effects	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>
Intercept	1.922	< 0.001***	−0.021	0.847	0.695	< 0.001***	20.841	< 0.001***	35.097	< 0.001***
Degree	0.022	0.473	0.072	0.031*	−0.003	0.935				
Clustering coefficient	0.126	0.052	0.069	0.048*	0.091	0.035*				
HIER							−2.426	0.5346	0.440	0.864
RCAVE							−2.640	0.4675	6.151	0.013*
RING							−7.950	0.0616	6.174	0.029*
No theory checks			−1.069	< 0.001***						
Pretest	0.002	0.223	0.002	0.500	−0.003	0.175				
Second round	0.275	< 0.001***	0.370	< 0.001***	−0.501	< 0.001***	14.888	< 0.001***	−3.218	0.160
Third round	0.294	< 0.001***	0.408	< 0.001***	−0.481	< 0.001***	7.716	< 0.001***	−5.005	0.032*
Factoid set 2	−0.014	0.772	−0.216	0.020*	0.227	0.067	−5.013	0.147	9.536	< 0.001***
Factoid set 3	0.011	0.791	0.128	0.175	0.176	0.17	4.275	0.229	3.051	0.193
<i>n</i>	816		816		510		51		51	
Model type	GLMM (ZIP)		LMM		LMM		OLS		OLS	
Unit of analysis	Individual		Individual		Individual		Aggregate		Aggregate	

Note. DV, dependent variable.

p* < 0.05; *p* < 0.01; ****p* < 0.001.

Table 4 Performance

DV	Pay per minute			
Model number	10		11	
Random effects	Variance			
Individual	7.240			
Run	8.231			
Fixed effects	Estimate	<i>p</i>	Estimate	<i>p</i>
Intercept	16.199	< 0.001***	392.14	< 0.001***
<i>Degree</i>	1.441	0.003**		
<i>Clustering coefficient</i>	1.364	0.216		
<i>HIER</i>			−26.24	0.172
<i>RCAVE</i>			−36.25	0.046*
<i>RING</i>			−65.73	0.002**
<i>Pretest</i>	0.055	0.022*		
<i>Second round</i>	7.505	< 0.001***	118.31	< 0.001***
<i>Third round</i>	10.689	< 0.001***	183.04	< 0.001***
<i>Factoid set 2</i>	−9.270	< 0.001***	−132.86	< 0.001***
<i>Factoid set 3</i>	−6.114	< 0.001***	−95.87	< 0.001***
<i>n</i>	816		51	
Model type	LMM		OLS	
Unit of analysis	Individual		Aggregate	

p* < 0.05; *p* < 0.01; ****p* < 0.001.

can lead people to accept others' answers rather than generate their own (Galbraith 1974, O'Reilly 1980, Schneider 1987, Speier et al. 1999), we found that degree (Model 6) was correlated with less exploration (greater copying). However, unlike clustering, greater degree was not correlated with any increased tendency to copy incorrect answers (Model 7) among those who copied at all.

Performance

Clustering led to better performance at the collective level, but individual subjects in more clustered positions did not perform better (see Table 4, Models 10 and 11). In other words, in the context of this experimental task, the networks that explored information space most also performed better. The fact that individuals in clustered positions did not perform better than others in their experimental run is consistent with the notion that more extensive exploration of information space occurs as a result of more coordination among network members, rather than more search by any single member.

Comparing results from Models 6 and 7 may also help explain why being in a clustered position did not confer performance benefits to individuals. There is a distinction between when exploration is likely and when it would be beneficial. Again, Model 6 shows that both high degree and clustering are correlated with less exploration of solution space, but Model 7 shows that clustering was specifically associated with convergence on incorrect answers while degree was not.

The rewired caveman network had the second-worst mean performance. The ring network had the worst performance except in the first round of play, in which the rewired caveman had the worst (see Table 5, Models 12–15). As in Model 3 of Table 2, the hierarchy network's performance was not significantly different from that of the rewired caveman.

Discussion

One can frequently hear comments marveling about how small our world has become. With the advent and

Table 5 Pay per Minute by Network Structure

Treatment	CAVE		HIER		RCAVE		RING	
Model number	12		13		14		15	
Random effects	Variance		Variance		Variance		Variance	
Individual	9.003		6.950		3.774		4.318	
Run	11.201		3.1480		0.000		9.601	
Fixed effects	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>	Estimate	<i>p</i>
Intercept	15.823	< 0.001***	22.529	< 0.001***	12.385	< 0.001***	10.908	0.002**
Total facts received	5.004	0.010*	2.965	0.161	7.012	0.002**	13.818	< 0.001***
Pretest	0.144	< 0.001***	−0.060	0.177	0.004	0.918	0.142	0.009**
Second round	9.868	< 0.001***	5.653	0.007**	7.967	< 0.001***	1.105	0.716
Third round	8.851	< 0.001***	11.723	< 0.001***	12.809	< 0.001***	3.489	0.357
Factoid set 2	−12.139	< 0.001***	−6.556	0.001**	−5.343	< 0.001***	−6.371	0.066
Factoid set 3	−5.978	0.015*	−8.975	< 0.001***	−2.083	0.245	−9.662	0.001**
<i>n</i>	272		176		224		144	
Unique individuals	154		150		135		136	
Runs	17		11		14		9	
Model type	LMM		LMM		LMM		LMM	
Unit of analysis	Individual		Individual		Individual		Individual	

p* < 0.05; *p* < 0.01; ****p* < 0.001.

accelerating adoption of increasingly powerful communication technologies and more global enterprises, our world is becoming ever more interconnected at every scale. In network terms, small-world networks have long been associated with surprisingly extensive diffusion of a given piece of information (Travers and Milgram 1969, Granovetter 1973, Watts and Strogatz 1998), in part as a result of the existence of short paths between any pair of individuals. It is therefore surprising, if not quite contradictory to those findings, that our subjects were less interdependent in the rewired caveman treatment—a small-world network—than in the other networks. Collaboration between network members ought to result in their performances being correlated, but in the rewired caveman, there was zero correlation of performance of individuals within the same run, suggesting less—or less effective—collaboration by experimental subjects. The small-world runs also had the least sharing and a high redundancy of facts found by search. In sum, the overall character of results from these runs was that people were more on their own than in the other networks.

But the existence of short paths between any given pair of individuals is not the only impact of greater communications connectivity. Broadcasts, publications, all manner of digital information systems (such as social media, topic-specific RSS feeds, and mobile applications), and even (at the global scale) air travel and international trade all promote mutual awareness and shared knowledge, much as clusters do in smaller-scale network terms. Our results can be understood to tell us a little more about the effects of that greater connectivity. We have argued throughout that the key feature of clustering is mutual awareness: within a cluster, everybody is aware of what everybody else is doing. In information space, this promotes exploration by allowing a sort of emergent coordination to occur inasmuch as people tend to avoid duplicating work they know has already been done. In solution space, it inhibits exploration by allowing more rapid convergence on a consensus about the solutions to problems.

The more connected we are, the more coordinated we become—either in a self-organized, emergent, “invisible hand” sort of way or in an intentional delegation sort of way—and the greater the diversity of what we do and can find out. We can celebrate how improved connectivity is making us ever better at coordinating our exploration of the facts of the world, just as the 16-person networks did in this experiment. Exploration for facts is becoming increasingly easier as global networks become increasingly dense. Greater interconnectivity can promote other emergent forms of coordination at the global scale as well. Increased geographic division of labor and the creation of niche communities of interest that could not otherwise sustain themselves are examples of increased diversity as a result of greater

interconnectivity—or, one could say, increased coordination and aggregate exploration of information space.

At the same time, the results above should make us extremely cautious about what that clustering means for our interpretation of this bounty of facts and how broadly we explore the answers that might be drawn from it. And with respect to the increased diversity of actions undertaken by the aggregate of humanity, we might worry that the way we understand those actions is becoming increasingly similar. Although we have more different types of goods and services than ever before, we have little diversity in economic policies, which is why we hear warnings of a spreading global monoculture and “McDonaldization.”

For knowledge-intensive organizations, the implication is that connecting everybody with increasingly high-bandwidth communications technologies may improve coordination but reduce diversity in the knowledge created within the firm (Benner and Tushman 2003). One possibility is that organizations could adopt different communications structures for different phases of collective problem solving. When information gathering and sharing is important, clustering will aid in greater exploration inasmuch as the information is not yet interpreted. Ordinarily, however, exploration of information space is guided by hypotheses or mental models (whether explicit or tacit). Our notions of the information space most relevant to explore derive from our working theories about the world. If a team wishes to find a better protective coating for an electronics product, it is very unlikely that any individual team member will go looking for information about the sugar content of fruit, no matter how well coordinated the members are in collectively avoiding duplicating their information-gathering work, because it is very unlikely that anyone’s working theory of protective coatings requires such information. In other words, it is inevitable that some degree of interpretation is always occurring in the minds of the information gatherers. If we wish to encourage the widest possible exploration of *relevant* information space, individuals should be arranged in clusters and extensively share their raw information with their neighbors, but they should keep the working theories that guided their exploration to themselves.

Prior literature shows that once information is in hand and diverse interpretations of that information are desired for generating theories or solutions, then less clustering is desirable, even within organizational subgroups, so that individuals do not prematurely coalesce on a consensus. In our experimental task, this was not a driver of net performance, but we note that it could well be crucial for other problems.

Another organizational response would be to design communications infrastructures that could somehow separate facts from figuring and adopt differently structured communication networks for each category. In other

words, rather than allow the march of technology to dictate organizational performance, it is possible to imagine technology being harnessed to achieve different performance goals. Even without the separation of facts and figuring, the results of this study are likely to be especially relevant for computer-mediated problem solving because of the ease of manipulating the structure in which participants communicate. Internal social networking (Leonardi 2014), knowledge management software, distributed teams (Mortensen and Neeley 2012), and external crowdsourcing platforms seem to be fertile grounds for testing these implications.

Future Extensions

Future Experiments. Building on this basic finding, much work could be done to refine the theory developed here and establish boundary conditions. For example, further experiments should investigate whether these results hold across a wider range of network structures and when information and solution spaces are much more rugged than those used here. Additionally, several findings that were statistically significant at the individual level were not significant at the collective level. Although clustering was associated with less redundancy of facts found, more theory checking, and more copying of incorrect theories, we did not find statistically significant differences between treatments at the collective level with respect to these outcomes. This is probably due to a dramatically lower number of person-run observations (51 versus 816), but other factors might also be at work.

One interesting pattern in our findings was that the results for the rewired caveman and ring networks were similar on certain outcome variables, despite those network treatments being so dissimilar in degree and centralization. Future work should focus on these aspects of structure to better tease out how they relate to the present results on clustering.

Replication Beyond Experiments. Empirical work should also ask whether these results hold when both imposed structure and social capital are operating to shape knowledge networks, which was not the case in our experiment. We chose to pursue an experiment in the laboratory because it allowed us to impose an interaction structure on a set of human beings and study its effects on different aspects of problem-solving behavior and performance. Because we used the experimental method, we were able to isolate and identify the complex fundamental effects of the interaction structure itself, something that observational studies of knowledge in social networks, with their richness and multiple layers, cannot do.

Nonetheless, the price we pay for greater certainty in terms of internal validity is less certainty in terms

of external validity. In real social networks, the correlation between clustering and shared mutual knowledge has been one of the primary reasons for studying clustering in the first place; important papers in this literature have sought to theorize the nature and consequences of that relationship (e.g., Granovetter 1973, Hansen 1999, Reagans and Zuckerman 2001, Sparrowe et al. 2001, Cummings and Cross 2003, Reagans and McEvily 2003, Burt 2004, Reagans et al. 2004). In the lab, the usual dependent variable—shared knowledge as it exists in vivo—cannot be present, nor can the usual independent variable, social network structure as it has been operationalized with persistent affective ties based on common experience or shared identity.

We do not mistakenly take the notion of an experimentally imposed network structure, defined as the pattern of communication ties, to be equivalent to an emergent “social structure” or “organization structure” in general. In the real world, an individual rarely has the agency to impose a network structure on an organization the way we did in the lab. We study if and how communication network structure constrains and creates opportunities for problem-solving behaviors and performance, even without the influence of other aspects of social or organization structure. We unfortunately cannot know how additional layers of network structure (such as trust, affect, and common experience) might bear on the results of this experiment. Our intention is to leave strong evidence about the differential impact of clustering on exploration for facts and exploration for solutions, opening the door for further research to take our findings back to the rich observational literature in this field.

Future Observational Work. Observational work could ask whether these results hold when communication over network ties is more costly or noisy. In our experiment, the average subject used all available network ties 93% of the time when sharing information. If communication were more costly, this would probably not be the case, though it is unclear how the results would change (Ghosh and Rosenkopf 2015). Additionally, our subjects shared information rapidly over a short time scale; different outcomes might result from slower sharing over a long time scale. Moreover, both experimental and observational studies could be enriched by the inclusion of embedded or tacit forms of knowledge (e.g., Argote and Miron-Spektor 2011), information accuracy (Kang et al. 2014), organizational culture considerations (e.g., Chatman and Barsade 1995, Chatman et al. 1998), network externalities (Levine and Kurzban 2006), and identity effects (e.g., Phillips et al. 2004, Liu and Srivastava 2015) as moderators of the effects we found. As for application of our results in an organization, the presence of social capital and shared tacit knowledge raises the question of how much power a “network designer” has to stipulate network structure when it cannot be imposed

through a computer interface. Organizations that wish to take advantage of our results and those from other experiments must combine these results with research on how networks come to have the structures that they do—through a combination of exogenous and endogenous factors (e.g., Chown and Liu 2015). Still, as we note above, our results do suggest applications for organizational information systems to mediate communication networks in ways that are specific to exploration for information versus exploration for solutions.

Implications: The Trade-off Between Facts and Figuring

Although it is well established that network structure can influence problem-solving performance, a clear understanding of the role of clustering, a basic structural network variable, has remained elusive. By theoretically and experimentally disentangling exploration for information from exploration for solutions—two core but separate domains of problem solving—we interpret the results of this study to indicate that clustering has opposite effects in those two domains. It promotes exploration through information space but depresses exploration through solution space. Whether increased clustering improves or impairs performance will therefore depend on whether the immediate task or problem-solving stage benefits more from exploration of facts or from the figuring that comes through the exploration of theories that interpret those facts.

According to the theoretical framework in prior work that only considers exploration of a single (solution) space, one would not know whether information search was an example of exploration or exploitation until it was revealed to be in the service of refining an existing solution or discovering a new one. The single-space framework does not, therefore, identify any relationship between network clustering and information search.

By considering the “search” in information search to be an example of exploration of a different space, we make it clear that clustering does have predictable effects on information search. To fully understand the role of network clustering in problem solving, one must consider the effects of clustering on exploration of both spaces, as well as the relative importance of exploration for information versus exploration for solutions in a given problem setting.

Awareness of the differential performance effects of clustering for problem solving in information space and in solution space raises two challenges. For networks of people trying to solve problems—whether they represent groups, organizational units, whole organizations, or clusters of organizations—the challenge is one of leadership. Leaders need to find ways to pair the domain of the problem-solving task—either facts or figuring—with an appropriate network structure, whether clustered or not, to improve problem-solving performance. For scholars

of networks and of information science, the challenge is one of further research. Integrating our basic finding of the distinction between facts and figuring into the examination of how different network structures affect performance may not only help resolve existing conflicts in disparate yet interconnected literatures but also open up substantial opportunities for a more coherent understanding of how we can set the conditions for problem-solving success in networks.

Clustering is a double-edged sword. It can encourage members of a network to generate more nonredundant information, but it can also discourage theoretical exploration. Until one knows whether a problem-solving task involves searching for facts or searching for answers, it is impossible to predict the influence of clustering on organizational performance.

Supplemental Material

Supplemental material to this paper is available at <http://dx.doi.org/10.1287/orsc.2015.0980>.

Acknowledgments

The authors extend their sincere thanks to Ray Reagans and two anonymous reviewers for their insights and helpful feedback. Thanks are also due to Allan Friedman, David Alberts, Mary Ruddy, and conference participants at the Workshop on Information in Networks, Organization Science Winter Conference, Workshop on Information Systems and Economics, INGRoup, Collective Intelligence, and the Harvard Work, Organizations, and Markets Seminar. Finally, the authors thank the organizing committee at INGRoup for selecting this paper for the 2014 Outstanding Paper Award. All errors are the authors' own. Screenshots used with permission from Azigo Inc. This work was supported by United States Department of Defense, Command and Control Research Program [Grant W74V8H-06-D-0008] and Army Research Laboratory [Grant ARL NS-CTA W911NF-09-2-0053]. Any opinions, findings, conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the funding agencies, or the U.S. government.

References

- Anjos F, Reagans R (2013) Commitment, learning, and alliance performance: A formal analysis using an agent-based network formation model. *J. Math. Sociol.* 37(1):1–23.
- Aral S, Van Alstyne M (2011) The diversity-bandwidth trade-off. *Amer. J. Sociol.* 117(1):90–171.
- Argote L, Miron-Spektor E (2011) Organizational learning: From experience to knowledge. *Organ. Sci.* 22(5):1123–1137.
- Bates D, Maechler M, Bolker B (2012) lme4: Linear mixed-effects models using Eigen and R syntax. R package version 0.999999-0, <http://CRAN.R-project.org/package=lme4>.
- Bavelas A (1950) Communication patterns in task-oriented groups. *J. Acoustic. Soc. America* 22(6):725–730.
- Benner MJ, Tushman ML (2003) Exploitation, exploration, and process management: The productivity dilemma revisited. *Acad. Management Rev.* 28(2):238–256.

- Bernstein ES (2012) The transparency paradox a role for privacy in organizational learning and operational control. *Admin. Sci. Quart.* 57(2):181–216.
- Burt RS (2004) Structural holes and good ideas. *Amer. J. Sociol.* 110(2):349–399.
- Centola D (2010) The spread of behavior in an online social network experiment. *Science* 329(5996):1194–1197.
- Centola D, Macy M (2007) Complex contagions and the weakness of long ties. *Amer. J. Sociol.* 113(3):702–734.
- Chatman JA, Barsade SA (1995) Personality, organizational culture, and cooperation: Evidence from a business simulation. *Admin. Sci. Quart.* 40(3):423–443.
- Chatman JA, Polzer JT, Barsade SG, Neale MA (1998) Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Admin. Sci. Quart.* 43(4):749–780.
- Chown JD, Liu CC (2015) Geography and power in an organizational forum: Evidence from the U.S. Senate Chamber. *Strategic Management J.* 36(2):177–196.
- Christensen CM, Carlile PR (2009) Course research: Using the case method to build and teach management theory. *Acad. Management Learn. Ed.* 8(2):240–251.
- Cummings JN, Cross R (2003) Structural properties of work groups and their consequences for performance. *Soc. Networks* 25(3):197–210.
- Dennis AR (1996) Information exchange and use in group decision making: You can lead a group to information, but you can't make it think. *MIS Quart.* 20(4):433–457.
- Enemark D, McCubbins MD, Weller N (2014) Knowledge and networks: An experimental test of how network knowledge affects coordination. *Soc. Networks* 36(January):122–133.
- Galbraith JR (1974) Organization design: An information processing view. *Interfaces* 4(3):28–36.
- Ghosh A, Rosenkopf L (2015) Shrouded in structure: Challenges and opportunities for a friction-based view of network research. *Organ. Sci.* 26(2):622–631.
- Granovetter MS (1973) The strength of weak ties. *Amer. J. Sociol.* 78(6):1360–1380.
- Guetzkow H, Simon HA (1955) The impact of certain communication nets upon organization and performance in task-oriented groups. *Management Sci.* 1(3–4):233–250.
- Hadfield JD (2010) MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *J. Statist. Software* 33(2):1–22.
- Hansen MT (1999) The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Admin. Sci. Quart.* 44(1):82–111.
- 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.
- Kang R, Kane AA, Kiesler S (2014) Teammate inaccuracy blindness: When information sharing tools hinder collaborative analysis. *Proc. Annual Meeting Comput. Supported Collaborative Work (CSCW '14)* (ACM, New York), 797–806.
- Kearns M, Suri S, Montfort N (2006) An experimental study of the coloring problem on human subject networks. *Science* 313(5788):824–827.
- Kearns M, Judd S, Tan J, Wortman J (2009) Behavioral experiments on biased voting in networks. *Proc. Natl. Acad. Sci. USA* 106(5):1347–1352.
- Kuhn F, Wattenhofer R (2006) On the complexity of distributed graph coloring. *Proc. 25th Annual ACM Sympos. Principles Distributed Comput. (PODC '06)* (ACM, New York), 7–15.
- Lazer DL, Friedman A (2007) The network structure of exploration and exploitation. *Admin. Sci. Quart.* 52(4):667–694.
- Leavitt HJ (1951) Some effects of certain communication patterns on group performance. *J. Abnormal Soc. Psych.* 46(1):38–50.
- Leonardi PM (2014) Social media, knowledge sharing, and innovation: Toward a theory of communication visibility. *Inform. Systems Res.* 25(4):796–816.
- Levine SS, Kurzban R (2006) Explaining clustering in social networks: Towards an evolutionary theory of cascading benefits. *Managerial Decision Econom.* 27(2–3):173–187.
- Lewis M (2010) *The Big Short: Inside the Doomsday Machine* (W. W. Norton & Company, New York).
- Liu CC, Srivastava SB (2015) Pulling closer and moving apart: Interaction, identity, and influence in the U.S. Senate, 1973 to 2009. *Amer. Sociol. Rev.* 80(1):192–217.
- March JG (1991) Exploration and exploitation in organizational learning. *Organ. Sci.* 2(1):71–87.
- Mason WA, Watts DJ (2012) Collaborative learning in networks. *Proc. Natl. Acad. Sci. USA* 109(3):764–769.
- Mason WA, Jones A, Goldstone RL (2008) Propagation of innovations in networked groups. *J. Experiment. Psych.: General* 137(3):422–433.
- McCubbins MD, Paturi R, Weller N (2009) Connected coordination network structure and group coordination. *Amer. Politics Res.* 37(5):899–920.
- Mortensen M, Neeley TB (2012) Reflected knowledge and trust in global collaboration. *Management Sci.* 58(12):2207–2224.
- Mulder M (1960) Communication structure, decision structure and group performance. *Sociometry* 23(1):1–14.
- O'Reilly CA III (1980) Individuals and information overload in organizations: Is more necessarily better? *Acad. Management J.* 23(4):684–696.
- Phillips KW, Mannix EA, Neale MA, Gruenfeld DH (2004) Diverse groups and information sharing: The effects of congruent ties. *J. Experiment. Soc. Psych.* 40(4):497–510.
- R Core Team (2012) *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna). <http://www.R-project.org/>.
- Reagans R, McEvily B (2003) Network structure and knowledge transfer: The effects of cohesion and range. *Admin. Sci. Quart.* 48(2):240–267.
- Reagans R, Zuckerman EW (2001) Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organ. Sci.* 12(4):502–517.
- Reagans R, Zuckerman EW, McEvily B (2004) How to make the team: Social networks vs. demography as criteria for designing effective teams. *Admin. Sci. Quart.* 49(1):101–133.
- Schneider SC (1987) Information overload: Causes and consequences. *Human Systems Management* 7(2):143–153.
- Shaw ME (1954) Some effects of problem complexity upon problem solution efficiency in different communication nets. *J. Experiment. Psych.* 48(3):211–217.
- Sparow RT, Liden RC, Wayne SJ, Kraimer ML (2001) Social networks and the performance of individuals and groups. *Acad. Management J.* 44(2):316–325.
- Speier C, Valacich JS, Vessey I (1999) The influence of task interruption on individual decision making: An information overload perspective. *Decision Sci.* 30(2):337–360.
- Stasser G, Stewart D (1992) Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *J. Personality Soc. Psych.* 63(3):426–434.
- Stasser G, Titus W (2003) Hidden profiles: A brief history. *Psych. Inquiry* 14(3–4):304–313.
- Travers J, Milgram S (1969) An experimental study of the small world problem. *Sociometry* 32(4):425–443.
- Watts DJ (1999) Networks, dynamics, and the small-world phenomenon. *Amer. J. Sociol.* 105(2):493–527.

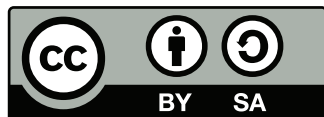
Watts DJ, Strogatz SH (1998) Collective dynamics of “small-world” networks. *Nature* 393(6684):440–442.

Jesse Shore is assistant professor of information systems at the Boston University Questrom School of Business. He received his Ph.D. from the Massachusetts Institute of Technology. His research interests include technological platforms for communication and collaboration and the consequences of social and communication network structure on the creation and transfer of knowledge.

Ethan Bernstein is assistant professor in the Organizational Behavior Unit at Harvard Business School. He received his

doctorate from Harvard University. His research is focused on the relationship between transparency/privacy and productivity in organizations. He seeks to understand how observability affects learning, innovation, and performance—for both the observer and the observed—in increasingly transparent workplaces.

David Lazer is Distinguished Professor of Political Science and Computer and Information Science, Northeastern University, and visiting scholar, Harvard University. He received his Ph.D. from the University of Michigan. His research focuses on the nexus of network science, computational social science, and collaborative intelligence. He is the founder of the citizen science behavioral experimental website VolunteerScience.com.



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License. You are free to download this work and share with others for any purpose, even commercially if you distribute your contributions under the same license as the original, and you must attribute this work as “*Organization Science*. Copyright 2015 INFORMS. <http://dx.doi.org/10.1287/orsc.2015.0980>, used under a Creative Commons Attribution License: <http://creativecommons.org/licenses/by-sa/4.0/>.”