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

Citation

Published Version

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:12111354

Terms of Use

Share Your Story

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

Jesse Shore
Ethan Bernstein
David Lazer

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

ABSTRACT

Using data from a novel laboratory experiment on complex problem solving in which we varied the network structure of 16-person organizations, we investigate how an organization's network structure shapes performance in problem-solving tasks. Problem solving, we argue, involves both *search for information* and *search for solutions*. Our results show that the effect of network clustering is opposite for these two important and complementary forms of search. Dense clustering encourages members of a network to generate more diverse information, but discourages them from generating diverse theories: in the language of March (1991), 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 were able to 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 the challenges inherent in manipulating network structure to enhance performance, as the communication structure that helps one antecedent of successful problem solving may harm the other.

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

Introduction

How does the clustering of organizational networks affect problem-solving behavior? Unfortunately, answers to that question remain incomplete. Clustering is the degree to which the people with whom a person is connected are themselves connected to each other. Substantial recent research would imply that clustering can improve problem-solving performance by increasing coordination (e.g., Kearns, Suri and Montfort, 2006; McCubbins et al, 2009), supporting the managerial trend towards efforts to increase connectedness in workplaces worldwide. In contrast, equally powerful recent research suggests clustering

can undermine performance by fostering an unproductive imbalance between exploration and exploitation (Lazer and Friedman, 2007; Mason, Jones, and Goldstone, 2005; Mason and Watts, 2012). A generalized net effect of clustering on problem-solving performance remains unresolved.

We present new evidence from a laboratory experiment that brings these disparate findings together under a single theoretical roof, bringing us a step closer to resolution of that question. We hypothesize that problem solving requires both *searching for information* (for the facts that may be important pieces of the puzzle) and *searching for solutions* (for theories that combine puzzle pieces into an answer). Our reading of existing literature suggests to us that the types of communication network structures that support searching for information may be different from those that support searching for solutions. We therefore adopt a novel, data-rich experimental platform that emphasizes verisimilitude with collective problem-solving tasks that people might confront in real organizational settings, incorporating search and sharing of both information and solutions.

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 be aware of each other's theories. This mutual knowledge facilitates an efficient search for additional information. However, the mutual awareness of each other's theories results in a convergence in interpreting that information, resulting in less exploration of theory space.

The same network structure, therefore, can 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 process of collective problem solving: structures that are good for performance now may be bad for performance later. We conclude with a discussion of implications for both theory and practice and suggest directions for future research.

Information and Solution Spaces
Overview: Clustering, Mutual Knowledge, and Performance

Individuals in cohesive clusters accrue shared, mutual knowledge as a consequence of communicating with each other (Grannovetter, 1973; Hansen, 1999; Burt, 2004, as well as many others). How clustering impacts problem-solving performance, however, remains an open question. Although a hot topic that has attracted both attention and high-quality scholarship, inconsistent results make conclusions elusive.

For bodies of understanding in which such inconsistency persists, a finer categorization of the observed phenomenon can help to resolve contradictions and enable progress (Carlile and Christensen, 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 a search for information (facts) and a search for solutions (interpretations of facts, or theories). We organize the following review of the literature accordingly, concluding with testable propositions.

Searching Solution Spaces: Clustering Undermines Performance?

Since March (1991), problem solving has been seen as analogous to 'exploration' of an unknown landscape. When a good solution is not known, one must explore the landscape in search of one. Since exploration is costly and uncertain, however, there is an incentive for risk-averse individuals to 'exploit' good solutions that have been discovered by others rather than strike out in a new direction (March, 1991:85). This exploitation can pose a problem at the collective level, because the greater the level of exploration across an organization, the more likely the organization is to discover a high quality solution somewhere on that landscape.

Using agent-based modeling, Lazer and Friedman (2007) show that individuals in highly connected networks (such that 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 are thus 'efficient' in their ability to facilitate diffusion of good solutions among network

3

members and outperform less connected networks in the short run, as individuals in less connected networks are less likely to be aware of good solutions found elsewhere in the network. The flip side of the inefficiency of less-connected networks in diffusing existing solutions is more exploration of the solution landscape by those who are not yet exploiting the current best solution, thereby bringing more potential solutions into the network. Inefficient networks, Lazer and Friedman (2007) find, eventually converged on superior solutions to those in efficient networks, because a greater region of the problem space was explored before agents converged on the best available solution.

In a large experiment on human subjects, Mason and Watts (2012) also find that networks that collectively explored more did better in the long run. However, they found very different structural correlates of exploration than those of Lazer and Friedman (2007). While Lazer and Friedman (2007) found that inefficient networks with a long average path length between any two individuals explored more, Mason and Watts (2012) found that inefficient networks—with, importantly, a high level of local clustering not present in Lazer and Friedman's (2007) networks—explored less.

The importance of the effects of local clustering in Mason and Watts (2012) can be explained by reference to work by Centola and Macy (2007; Centola, 2010), who find that adopting a neighbor's solution is more likely to occur within, rather than between, clusters of ties. Novel, exploratory solutions are both uncertain in advance and have material consequences. Given these conditions, the act of observing that multiple other people have adopted the same existing solution—as is more likely within clusters than between them—makes further adoption even more likely, because of the apparent emergence of a consensus opinion. In the language of March (1991), this is to say that exploration of solution space should be less extensive within clusters, and Mason and Watts (2012) showed strong evidence of this.

Searching Information Spaces: Clustering Aids Performance?

Recent Graph Coloring Experiments. There is an expanding experimental literature on collective problem solving in networks, primarily using variants of the "distributed graph coloring problem" as an experimental task (e.g., Kuhn and Wattenhofer, 2006; Kearns, Suri and Montfort, 2006; Kearns, et. Al,

2009; Judd, Kearns, and Vorobeychick, 2010). In graph coloring tasks, subjects must choose from among a discrete set of colors such that they do or do not match the choice of their neighbors. There is no subjective interpretation required for these tasks: each subject takes in information about their neighbors and selects their own color according to the instructions. The lack of interpretation makes these tasks quite different from those in the previous section. The critical variable for success in these tasks is instead coordination. For our purposes here, it is unfortunate that these high-quality graph coloring studies largely investigate the effects of density rather than clustering. While this means that these results cannot be assumed to apply exactly to clustering, density and clustering are related measures: clustering is the degree to which everyone within a local neighborhood is connected with each other, while density is the analogous property of the whole network.

In general, greater density of ties improves performance in these tasks (McCubbins, et al., 2009). A major mechanism in this increased performance within dense clusters appears to be greater mutual knowledge. Knowing what your neighbors' neighbors are doing dramatically eases the distributed graph coloring problem (Enemark, McCubbins, and Weller, 2012). 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 search of information spaces.

Earlier Experiments. The first laboratory-based social network experiments were from MIT's Small Group Network Laboratory (SGNL) by Bavelas (1950) and Leavitt (1951) on 5-person networks. These experiments appear to contradict the graph coloring literature in that they find that 5-person star graphs (centralized, unclustered networks) were more effective in cooperative problem solving than 5-person complete cliques (maximally clustered networks) (Bavelas, 1950; Leavitt, 1951). In similar experiments, Guetzkow and Simon (1955) subsequently showed that the star graph all but guarantees that at least one 'central' person receives all the information available in the network, which she or he can then disseminate to the rest of the organization members. Shaw (1954) went on to argue that "complex" (meaning more difficult) problems were better solved in "decentralized structures" (clusters), but Mulder (1960) later established that ultimately centralized (unclustered) networks performed better for both

5

simple and difficult problems, once a centralized and coordinated decision structure evolves within the experimentally imposed communication structure.

The protocol deployed in these experiments differed from more recent experiments in a critical way, however, in that each act of communication could only be directed to a single individual. Unlike recent experiments in which multiple network ties could be used simultaneously, in these early experiments, only one tie could be used at a time. 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 it is surely possible to construct a real-world collective problem-solving task like the Bavelas experiments, these seem much more the exception than the norm. As long as communications can be addressed to multiple individuals at a time—for example, because they are all sitting around a table in a meeting, or because they can broadcast their status to multiple individuals on a social media platform—then clustering would ensure that those individuals had full access to shared information, which is argued to be associated with better performance in both recent experiments and the early experimental work. In sum, although these earlier results appear to diverge from those of the recent graph coloring experiments referenced above, the differences are probably due to differences in the experimental protocols; 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

These streams of research on search for solutions and searching for information, taken together, provide 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 by Figure 1.

-------------------- Insert FIGURE 1 about here --------------------

6

When everybody knows what everybody else is doing, it allows them 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 non-complementary, duplicated work. Similarly, the more individuals are aware of each others' interpretations, as would be the case within a cluster (Centola and Macy, 2007), 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. While clustering does not guarantee either outcome, by promoting mutual awareness, it does make it easier for individuals to both coordinate and copy to avoid redundancy.

An alternative way to frame the existing literature, as suggested by Figure 1, is that different types of exploration are valuable within the different domains. In a search for information, coordinated exploration may be valuable because it avoids duplicated work. In a search for solutions, uncoordinated exploration may be valuable because it avoids copying. Unfortunately, clustering simultaneously impacts coordination in both domains.

To illustrate the difference between the effects of clustering in the two domains, consider the case of market analysts making forecasts. To make forecasts, a range of data must first be gathered—the unemployment rate, GDP growth, the availability of credit, costs of inputs to production and so on. If analysts have shared and mutual information, they will be less likely to collect data that is already known, and 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, this information also requires a great deal of interpretation to produce a forecast. The more analysts have shared and mutual knowledge, the more likely they are to be influenced by each others' 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 United States mortgage market was sustainable (Lewis, 2010). Mutual and shared knowledge inhibits exploration of solution space (analysts can be said to be 'solving' the puzzle of what the future state of a market will be in this scenario).

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

Data and Methods
The Experimental Platform

The Task. To instrument the connection between clustering, solution search, information search, and problem-solving performance, we aimed to develop an experimental platform with several key characteristics: (1) 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; (2) maximum accessibility, which required the task to be easily understandable and solvable with expertise commonly available in our subject pool; and (3) maximum instrumentation, which required that actions taken by the participants be captured as richly as possible in subsequently analyzable data.

Based on those criteria, we selected a whodunit protocol, much like a game of Clue® or Cluedo®, in which the task involved piecing together clues to "connect the dots." Our task therefore bore some resemblance to the common murder mystery protocol in group research (Stasser and Stewart, 1992:429; Stasser and Titus, 2003), but with several key differences to instrument our research question: (a) as clustering of communication ties would be a variable of interest (see treatments 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 search (as opposed to only discussion), our clues would not all be distributed to members of the organization, but instead would be accessible via a Google-style search--

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 was enabled by technology (e.g., Dennis, 1996), sharing of information was less limited by synchronous airtime in discussion (i.e., just because one person was sharing didn't mean that others couldn't do so simultaneously). 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 United States Department of Defense's Command and Control Research Program called ELICIT (Experimental Laboratory for Investigating Collaboration, Information-sharing, and Trust), which had many of the characteristics we sought. While we modified much of the platform, we agreed to keep 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 the murder in Clue®). More detail is available in online appendix A.

Specifically, participants were faced with four logically-independent sub-problems to solve: (1) who would carry out the attack (group involved); (2) what would the target of the attack be (e.g. an embassy, a church, etc.); (3) where would the attack take place (country); and (4) when would the attack take place (with four interdependent components--month, day, hour, and AM/PM). Each question and sub-question had a dedicated text box in which to register an answer. Participants had 25 minutes with which to try to solve the problem.

There was a discrete set of actions available to experimental subjects. They could search for new clues by entering a keyword into a search text box and clicking a button, at which point the tool would search for that keyword amongst the 69-83 facts contained in the pre-determined factoid set for that problem. They could share clues they already had with one or more of their neighbors 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 sub-problems. Finally, they could check their neighbors' registered solutions at any point and any number of times during the experiment.

Participants were rewarded 15 cents per minute per sub-problem (3.75 cents per minute for each component of "when"), for a maximum of 60 cents (a penny per second) for each minute that they had the correct answer registered. Participants therefore had a strong incentive to record and adjust their theories as soon as they had developed a theory about the answer. Although success required help from neighbors—through sharing of clues, annotation of clues, and viewing of neighbors' registered solutions—we elected to provide incentives exclusively at the individual level. Competition or group-level rewards would have introduced new interdependencies in 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 is also an important topic of study, but we limit ourselves to the primary phenomenon for this experiment. Since the information space was large and time was limited to only 25 minutes, it was virtually impossible for an individual to find all clues necessary to solve the problem alone, ensuring that individuals 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 their 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 experiment's conclusion. Each experimental run lasted 25 minutes. Participants were given two clues at the start of each round, and were allowed to search for clues once per minute. Each clue was only relevant to one sub-problem. Some clues contained useless information, and some contained misleading information. Subjects had to combine multiple clues to conclusively arrive at the correct answer. The number of clues required to conclusively arrive at the correct answer for a single sub-problem ranged from 2 to 10, with a median of 5 and a mean of 5.3.

-------------------- Insert FIGURE 2 about here --------------------

10

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

After the pre-test, 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 of the experiment, 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 multivariate analyses, we include a "Round" variable, corresponding to which time through the experiment the data is from.

At any given time, lab space allowed us to run a maximum of 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 terrorist groups, country names, and place names) were randomly permuted to reduce the risk of contamination between sessions.

Prior to a run beginning, 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 within the laboratory. Participants were assigned a pseudonym in order to further obscure their identities from other subjects, and pseudonyms were shuffled before each round.

There were 417 unique individuals, who played a total of 1120 person-rounds. Participants were recruited through the subject pool of a large university in the northeast United States. The mean self-reported math SAT score was 716, and the mean verbal score was 701, consistent with reported data from the university. The ratio of self-reported gender was approximately equal, with 49.5% male and 50.5% female. The present 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 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 so-called "caveman" network (Watts, 1999), containing four four-person cliques. The "hierarchy" is likewise composed of four such cliques, but arranged in a conventional centralized structure. The "rewired caveman" is a small world network, constructed by removing links from the caveman, then adding links that create shortcuts through the network. As a result, individuals in the rewired caveman are "closer together" topologically: the most distant pair of individuals is only three hops away, and the average distance between all pairs is shorter than the caveman and the hierarchy. The rewired caveman is also more centralized and less clustered than the caveman. Finally, there is the "ring" network, which is neither clustered nor centralized.

-------------------- Insert FIGURE 3 about here --------------------

-------------------- Insert TABLE 1 about here --------------------

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

Outcome Variables

We treat exploration and exploitation as mutually exclusive classifications of a single action, as has been customary in the past (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 is important in our measurement of these constructs, because in different cases, one or the other is easier to observe. For example, it is easier to measure the amount or extent of information gathering (exploration) than mental processing of information already held (exploitation). Additionally, as we are working at both the collective and individual level, we specify measurements of exploration (exploitation) at each level separately.

12

Exploration in Information Space. In our experiment, information space is explored by means of searching for facts. We measure 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 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 the 'redundancy' of facts found. The more unique facts and the lower the redundancy of facts found, the greater the exploration of information space.

At the individual level, we measure the total number of facts found by a subject's own search (the search interface did not return facts that an individual already possessed, so the total number of facts represents the extent of exploration in information space by search) and the redundancy of facts received from neighbors (the ratio of total facts received to unique facts received).

Exploration and Exploitation in Solution Space. We measure 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 focus on exploitation because it is easier to observe. Specifically, exploitation in solution space takes the form of checking and then copying a neighbor's theory. We have time-stamped records of every action undertaken during the experiment, so checking neighbors' theories can be directly measured from the data. We define copying to be when an individual checks their neighbors' answers, and then registers an answer they observed the next time they enter their own theories, provided this occurs within 10 minutes of observation of the neighbors.

Establishing uniqueness of theories required us to be able to consolidate answers like "power plant" and "powerplant" and "electric power plant" into one theory, which we did in two steps. First, automated pre-processing 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, provided that it did not involve the first letter of the word). Second, we used a human coder to remove more substantial typos (such as transposing letters or whole phonemes) as well as answers where the intent was clearly the same (for example, we considered "power plant" and "electric power plant" to be the same).

13

Performance. Given that clustering is expected to have both positive and negative effects, depending on the domain of reference, we also measure overall performance. Performance was measured in pay per minute received by individuals. When measuring performance of an entire network, we simply took the sum of the pay per minute for each individual member of the network.

Statistical Framework

Wherever possible, we consider both individual and collective-level correlations. At the collective level, we have 51 independent data points, each corresponding to a single run of the experiment. For these models, we use ordinary least squares (OLS). At the individual level, we have 816 observations with two types of interdependence. First, they are 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 include random effects for both run and unique individual, and estimate linear mixed models (abbreviated LMM in the tables of results).

For discrete outcome variables at the individual level, we employed mixed-effects Poisson (GLMM-Poisson) regression (Bates, Maechler, and Bolker, 2012). The number of times a participant checked the answers of their neighbors exhibited zero-inflation, and thus a zero-inflated Poisson mixed-effects model (GLMM-ZIP) was estimated in a Markov Chain Monte Carlo framework (Hadfield, 2010). Statistical analysis was carried out in R (R Core Team, 2012). Analytical code and data will be made available at Dataverse (thedata.org).

The tables of results that we present in the following section contain evidence to support causal and non-causal inferences. The experimental treatments we impose 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 is lower than performance in the hierarchy, we can conclude that it is because of 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, like degree and clustering -- should be interpreted as correlations, rather than evidence of causality. Outcomes at the individual level depend not

14

only on the local structure, but also the structure of the remainder of the network, and node-level metrics do not account for this. Nonetheless, they can help us to understand the causal results at the collective level by providing more understanding of the behavior of individuals.

The result section below reports on each of the three sets of variables above in turn: 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.

Exploration

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, information. By being in a cluster, individuals tended to contribute more to the *collective* exploration through information space, not from more search, but instead 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 that they perform (Table 2, Model 1). Likewise, at the collective level, the clustered networks did not search at a different rate than the rewired caveman (Model 2 – note that the caveman network serves as the reference category).

-------------------- Insert TABLE 2 about here --------------------

15

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 network and the rewired caveman treatments (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 was also lower than the ring network and the rewired caveman treatments, although the difference was not significant -- a fact that we consider further below in the section on centralization. In addition to searching for information, exploration of information space at the collective level requires that information be effectively transferred throughout the network, and 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 decreased number of 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 copying of neighbors' answers (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 total number of unique theories registered at the whole organization level. Checking and copying the theories of neighbors both indicate less extensive exploration of solution space, and we find that clustering at the individual level is associated with more checks of neighbors' theories (Model 5: marginally significant finding) 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 of their neighbors (Model 7). 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

16

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 can lead individuals to accept others' answers rather than generate their own (Galbraith, 1974; O'Reilly, 1980; Schneider, 1987; Speier, Valacich, and Vessey, 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 that copied at all.

-------------------- Insert TABLE 3 about here --------------------

Performance. Clustering led to greater performance at the collective level, but at the individual level, subjects in more clustered positions did not perform better (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 to occur, and when exploration 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 had a mean performance greater only than the ring network, which had the worst performance except in the first round of play, in which the rewired caveman had the lowest mean performance of all the treatments, including the ring network (Table 5, Models 12-15). Like Model 3 above, the hierarchy was not significantly different from the rewired cave in terms of performance.

-------------------- Insert TABLE 4 about here --------------------

-------------------- Insert TABLE 5 about here --------------------

17

Discussion

One can frequently hear marveling comments about how small our world has become. With the advent and 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 been long associated with surprisingly extensive diffusion of a given piece of information (Travers and Milgram, 1969; Granovetter, 1973; Watts and Strogatz, 1998), due in part to the existence of short paths between any pair of individuals. It is therefore surprising, if not quite contradictory to those findings, that in the rewired caveman treatment -- a small world network -- people depended less on each other than in other network environments. Collaboration between network members ought to result in their performances being correlated, but in the rewired caveman, correlation of performance of individuals within the same run was zero, suggesting less or less effective collaboration by experimental subjects. The small world runs also had the least amount of 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 short paths between any given pair of individuals are not the only impact of greater communications connectivity. Broadcasts and publications, all manner of digital information systems including, social media, topic-specific RSS feeds, and mobile applications -- 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 in that people tended to avoid duplicating the work that they knew had already been done. In solution space, it inhibits exploration by allowing more rapid convergence on a consensus about the problems' solutions.

18

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 our 16-person networks did in this experiment. Searching for facts is becoming easier and 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 due to greater interconnectivity -- or, one could say, increased aggregate exploration of activity space.

At the same time, the results above should make us extremely cautious about what that clustering means for our interpretation of those now plentiful facts and how broadly we explore the possible answers from these facts that we are increasingly good at finding. 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, just as we are sometimes also warned of 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, then clustering is beneficial. For such a phase of information gathering and dissemination, clustering will aid in greater exploration in as much as the information is not yet interpreted. Ordinarily, however, searches of information space are guided by some sort of hypotheses or mental models (whether explicit or tacit). Our notions of the relevant information space to explore derive from our working theories about the world. If a team wishes to find a better protective coating for an electronic product, individual team members will probably not elect to

search the space of information about the sugar content of fruits, no matter how well coordinated they are in avoiding duplicating their information gathering work, because it is very unlikely that a 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, extensively share their raw information with their neighbors, but keep their working theories that guided their searches to themselves.

When information is in hand and it is desirable to generate diverse interpretations of it to generate theories or solutions, then prior literature shows that less clustering is desirable, even within organizational subgroups, such 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 a crucial factor 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 how technology could be harnessed to achieve different performance goals. Even without separation of facts and figuring, these results are likely to be especially relevant for computer-mediated problem solving, because of the ease of manipulating the communication structure of participants. Internal social networking and knowledge management software and external crowdsourcing platforms seem to be fertile ground 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.

Replication Beyond Experiments. Empirical work should also ask if these results hold when both imposed structure and social capital are operating to shape knowledge networks, which our experiment does not do. 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 empirical 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 human social networks, the correlation between clustering and shared, mutual knowledge has been one of the primary reasons to study clustering in the first place, and important papers in this literature have sought to theorize the nature and consequences of that relationship (e.g., Grannovetter, 1973; Hansen, 1999; Sparrowe et al, 2001; Reagans and Zuckerman, 2001; Reagans and McEvily, 2003; Cummings and Cross, 2003; Burt, 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 mistake the notion of experimentally imposed network structure, defined as the pattern of communication ties, to be equivalent to 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

searching for facts and searching for solutions, opening the door for further research to take our findings back to the rich empirical literature in this field.

     Future Empirical Work. Empirical work could ask if these results hold when communication over network ties is more costly or noisy. In our experiment, the average subject used all of their available network ties 93% of the time when they shared information. If communication were more costly, this would probably not be the case, and it is unclear how the results would change under these circumstances. Moreover, both experimental and observational studies could be enriched by inclusion of embedded or tacit forms of knowledge (e.g. Argote and Miron-Spektor, 2011) and by inclusion of organizational culture considerations (e.g., Chatman and Barsade, 1995; Chatman et al, 1998) as a moderator of this effect. As for intentional application of these 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 by use of a computer interface. For organizations to take advantage of these results and results from other experiments, they will necessarily need to combine them with research on how networks come to have the structure that they do, through a combination of exogenous and endogenous factors (e.g. Chown and Liu, 2014). On the other hand, as we note above, the results do suggest applications for organizational information systems to mediate communication networks in ways that are specific to search for information versus solutions.

Implications: The Tradeoff Between Facts and Figuring

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

exploration of facts or, instead, the figuring that comes through the exploration of theories that interpret those facts.

Awareness of the differential performance effects of clustering for problem solving in information space from problem solving in solution space presents two challenges. For networks of problem-solving individuals, whether they represent groups, organizational units, whole organizations, or clusters of organizations, the challenge is one of leadership, such that leaders find ways of pairing the domain of the problem-solving task, whether facts or figuring, with an appropriate network structure, whether clustered or not, to improve problem-solving performance. For scholars of both networks and 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 impact performance may help not only to resolve existing conflicts in disparate yet interconnected literatures, but also open up substantial opportunities for greater, coherent understanding of how we can set the conditions for problem-solving success in networks.

Clustering is a double-edged sword. It has the power to bring members of a network to generate more non-redundant information, but it also has the power to 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.

REFERENCES

Argote, L. and E. Miron-Spektor. 2011. Organizational learning: From experience to knowledge. *Organization Science*, 22(5): 1123–1137

Bates, D., M. Maechler, and B. Bolker. 2012. lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4

Bavelas, A. 1950. Communication patterns in task-oriented groups. *The Journal of the Acoustical Society of America*, 22(6): 725-730.

Benner, M. J., & Tushman, M. L. (2003). Exploitation, exploration, and process management: The productivity dilemma revisited. *Academy of Management Review*, 28(2), 238-256.

Burt, R.S. 2004. Structural holes and good ideas. *American Journal of Sociology* 110(2): 349-399.

Chatman, J.A. & Barsade, S. A. 1995. Personality, organizational culture, and cooperation: Evidence from a business simulation. *Administrative Science Quarterly*, 40(3): 423-443.

Chatman, J. A., Polzer, J. T., Barsade, S. G., & Neale, M. A. 1998. Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly*, 43(4), 749-780.

Christensen, C. M. & Carlile, P. R. 2009. Course research: using the case method to build and teach management theory. *Academy of Management Learning & Education*, *8*(2), 240–251

Centola, D. 2010. The spread of behavior in an online social network experiment. *Science*, 329(5996): 1194-1197.

Centola, D., and M. Macy. 2007. Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3): 702-734.

Chown, Jillian and Christopher Liu. 2014. Geography and Power in an Organizational Forum: Evidence from the U.S. Senate Chamber. *Strategic Management Journal.* Forthcoming.

Cummings, J. N., and R. Cross. 2003. Structural properties of work groups and their consequences for performance. *Social Networks*, 25(3): 197-210.

Dennis, A. R. 1996. Information exchange and use in group decision making: You can lead a group to information, but you can't make it think. *MIS Quarterly*, 20: 433-457.

Enemark, D., M.D. McCubbins, R. Paturi, and N. Weller. 2011. Does more connectivity help groups to solve social problems? *In Proceedings of the ACM Conference of Electronic Commerce*, 21-27.

Enemark, D., M.D. McCubbins, and N. Weller. 2012. Knowledge and networks: An experimental test of how network knowledge affects coordination. *Social Networks*.

Galbraith, J. R. 1974. Organization design: An information processing view. *Interfaces*, 4(3): 28-36.

Granovetter, M. S. 1973. The strength of weak ties. *American Journal of Sociology*: 1360-1380.

Guetzkow, H., and H.A. Simon. 1955. The impact of certain communication nets upon organization and performance in task-oriented groups. *Management Science*, 1(3-4): 233-250.

Hadfield, Jarrod D. 2010. MCMC methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*, 33(2): 1-22.

Hansen, M. T. 1999. The search-transfer problem: The role of weak ties in sharing knowledge across organization subunits. *Administrative Science Quarterly*, 44(1): 82-111.

Judd, S., M. Kearns, and Y. Vorobeychik. 2010. Behavioral dynamics and influence in networked coloring and consensus. *Proceedings of the National Academy of Sciences*, 107(34): 14978-14982.

Kearns, M., S. Judd, J. Tan, and J.Wortman. 2009. Behavioral experiments on biased voting in networks. *Proceedings of the National Academy of Sciences*, 106(5): 1347-1352.

Kearns, M, S. Suri and N. Montfort. 2006. An experimental study of the coloring problem on human subject networks. *Science* 313: 824-827

Kuhn, F. and R. Wattenhofer. 2006. On the complexity of distributed graph coloring. In *Proceedings of the Twenty-Fifth annual ACM symposium on Principles of Distributed Computing*: 7-15.

Lazer, D.L. and A. Friedman. 2007. The network structure of exploration and exploitation. *Administrative Science Quarterly*, 52(4): 667-694.

Leavitt, H. J. 1951. Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, 46(1): 38.

Lewis, M. 2010. *The big short: Inside the doomsday machine*. WW Norton & Company.

March, J. G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2(1): 71-87.

Mason, W. A.,Jones, A., and R.L. Goldstone. 2008. Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3): 422.

Mason, W.A., and D.J. Watts. 2012. Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, 109(3): 764-769.

McCubbins, M. D., R. Paturi, R., and Weller, N. 2009. Connected coordination network structure and group coordination. *American Politics Research*, 37(5): 899-920.

Mulder, M. 1960. Communication structure, decision structure and group performance. *Sociometry*, 23(1): 1-14.

O'Reilly, C. A. 1980. Individuals and information overload in organizations: is more necessarily better? *Academy of Management Journal*, 23(4): 684-696.

R Core Team. 2012. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reagans, R., and B. McEvily. 2003. Network structure and knowledge transfer: The effects of cohesion and range. *Administrative Science Quarterly*, 48(2): 240-267.

Reagans, R., and E.W. Zuckerman. 2001. Networks, diversity, and productivity: The social capital of corporate R&D teams. *Organization Science*, 12(4): 502-517.

Schneider, S. C. 1987. Information overload: Causes and consequences. *Human Systems Management*, 7(2): 143-153.

Shalizi, C. R. and A.C. Thomas. 2011. Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods & Research*, 40(2): 211-239.

Shaw, Marvin E. 1954. Some effects of problem complexity upon problem solution efficiency in different communication nets. *Journal of Experimental Psychology*, 48(3): 211-217

Sparrowe, R. T., R. C. Liden, S.J. Wayne, and M.L. Kraimer. 2001. Social networks and the performance of individuals and groups. *Academy of Management Journal*, 44(2): 316-325.

Speier, C., J.S Valacich, and I. Vessey. 1999. The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30(2): 337-360.

Stasser, G., and D. Stewart. 1992. Discovery of hidden profiles by decision-making groups: Solving a problem versus making a judgment. *Journal of Personality and Social Psychology*, 63(3): 426.

Stasser, G and W. Titus. 2003. Hidden profiles: A brief history. *Psychological Inquiry*, 14(3-4): 304-313.

Travers, J. and S. Milgram. 1969. An experimental study of the small world problem. *Sociometry*, 425-443.

Watts, D.J. 1999. Networks, dynamics, and the small-world phenomenon. *American Journal of Sociology*, 105: 493–527.

Watts, D. J. and S.H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature*, 393(6684): 440-442.

FIGURES

Figure 1: Impact of Clustering on Performance by Domain

Figure 1. Impact of Clustering on Performance by Domain
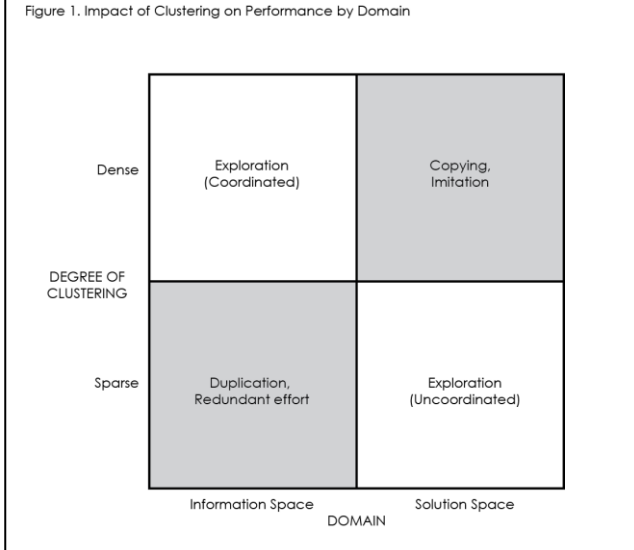
Figure 3: Network treatments used in the experiments.
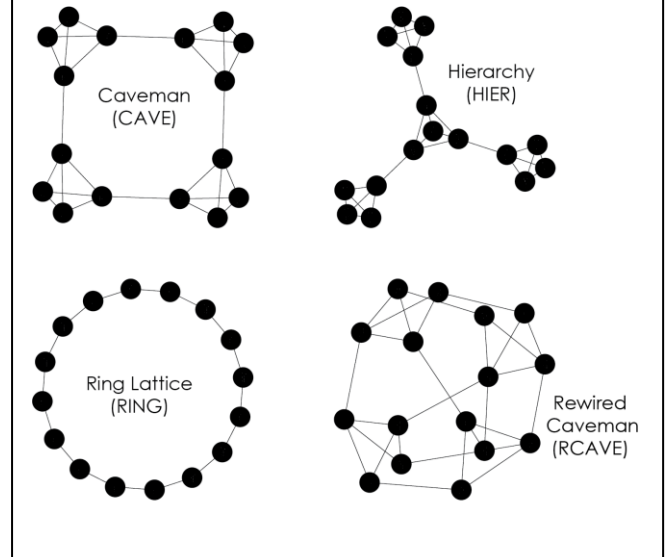
Figure 2: a screenshot of the web browser interface. Pictured is the "inbox" of an experimental subject.

TABLES

Table 1: *Descriptive statistics for networks*

	CAVE	RCAVE	HIER	RING
average degree	3.5	3.625	3.375	2
density	.467	.483	.45	.267
average path length	2.47	1.99	2.81	4
diameter	5	3	5	8
mean clustering coefficient	.667	.304	.727	0
centralization (eigenvector)	.033	.115	.161	0

Table 2: Exploration and Exploitation in Information Space

	Dependent variable							
	Total facts found by search				Total Searches / Unique Facts Found		Total receipts / Unique Facts Received	
Model number	1		2		3		4	
Random effects	Variance						Variance	
Individuals	0.012						0.001	
Run	0.000						0.006	
Fixed effects	Estimate	*p*	Estimate	*p*	Estimate	*p*	Estimate	*p*
(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***
Clust. Coef.	-0.019	0.423					-0.131	<0.001***
Hierarchy			2.424	0.722	0.114	0.198		
Rewired Cave			7.231	0.257	0.199	0.017*		
Ring Network			46.248	<0.001***	0.761	<0.001***		
Pre-test	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
n	816		816		51		816	
Model type	GLMM (Poisson)		OLS		OLS		LMM	
Unit of analysis	Individual		Aggregate		Aggregate		Individual	

Table 3: Exploration and Exploitation in Solutions Space

	Theory checking		Theory copying		Incorrect copying		Theory copying		Unique theories	
Model number	5		6		7		8		9	
Random effects	Variance		Variance							
Individuals	0.102		0.028		0.200					
Run	0.000		0.035		0.039					
Fixed effects	Estimate	p	Estimate	p	Estimate	p	Estimate	p		
(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				
Clust Coef	0.126	0.052.	0.069	0.048*	0.091	0.035*				
Hierarchy							-2.426	0.5346	0.440	0.864
Rewired Cave							-2.640	0.4675	6.151	0.013*
Ring Network							-7.950	0.0616.	6.174	0.029*
No theory checks			-1.069	<0.001***						
Pre test	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
n	816		816		510		51		51	
Model type	GLMM (ZIP)		LMM		LMM		OLS		OLS	
Unit of analysis	Individual		Individual		Individual		Aggregate		Aggregate	

Table 4: Performance

	Pay per minute			
Model number	10		11	
Random effects	Variance			
Individuals	7.240			
Run	8.231			
Fixed effects	Estimate	p	Estimate	p
(Intercept)	16.199	<0.001***	392.14	<0.001***
Degree	1.441	0.003**		
Clust. Coef.	1.364	0.216		
Hierarchy			-26.24	0.172
Rewired Cave			-36.25	0.046*
Ring Network			-65.73	0.002**
Pre-test	0.055	0.022*		
Second round	7.505	<0.001***	118.31	<0.001***
Third round	10.689	<0.001***	183.04	<0.001***
Factoid set 2	-9.270	<0.001***	-132.86	<0.001***
Factoid set 3	-6.114	<0.001***	-95.87	<0.001***
n	816		51	
Model type	LMM		OLS	
Unit of analysis	Individual		Aggregate	

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	
Individuals	9.003		6.950		3.774		4.318	
Run	11.201		3.1480		0.000		9.601	
Fixed effects	Estimate	*p*	Estimate	*p*	Estimate	*p*	Estimate	*p*
(Intercept)	15.823	<0.001***	22.529	<0.001***	12.385	<0.001***	10.9078	0.002**
Total Facts Rec'd	5.004	0.010*	2.965	0.161	7.012	0.002**	13.818	<0.001***
Pre-test	0.144	<0.001***	-0.060	0.177	0.004	0.918	0.1415	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	0.2449	-9.662	0.001**
n	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	

ONLINE APPENDIX A: Explanation of the ELICIT Experimental Tool

This appendix provides an overview of the experimental tool used for this research. For a more detailed introduction, a full lab protocol and instructional video is available from the authors upon request. The main part of the subject's screen is always the inbox. The inbox will show all clues which a participant has received from any source. A participant can receive clues from three sources: "New Data" from the system (at the start of the experiment), a "Search Result" from a web search, and clues shared by other participants.

At the beginning of any session, the inbox will contain a message from the system moderator assigning the subject a pseudonym and two initial clues, as per Figure A1.

Figure A1: Inbox at the Beginning of a Session

From the inbox screen, a participant can engage in one of several actions, including three main ones: (a) share a clue; (b) search for a clue; and (c) identify an answer.

a) **Share**: At any time during the experiment, a subject may choose to share any clue in the inbox with certain other subjects in the experiment, just as one might forward an email. To do so, the subject simply (1) selects the radio button next to the relevant clue; (2) clicks on "Share"; (3) on the pop-up window, selects the other subjects who will receive the clue; and (4) annotates the clue, if desired (see Figure A2).

The clue will then appear in the recipients' inboxes. As an example, in the inbox in Figure A3, Quinn shared two annotated clues with Pat. Note that Pat can respond to each of them by clicking "Reply".


Figure A2: How to Share

31
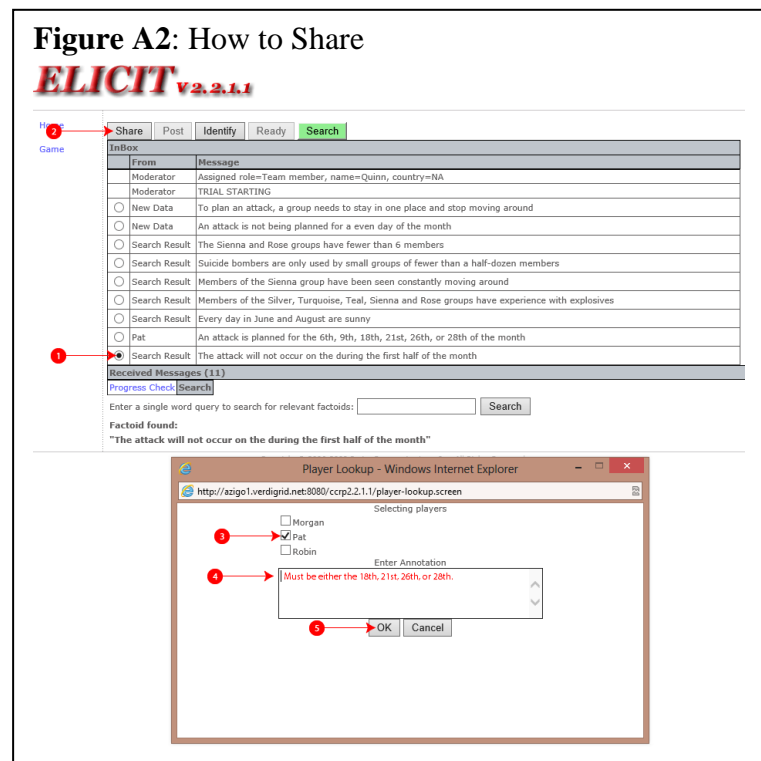
Figure A3: Example of Annotated, Shared Clues
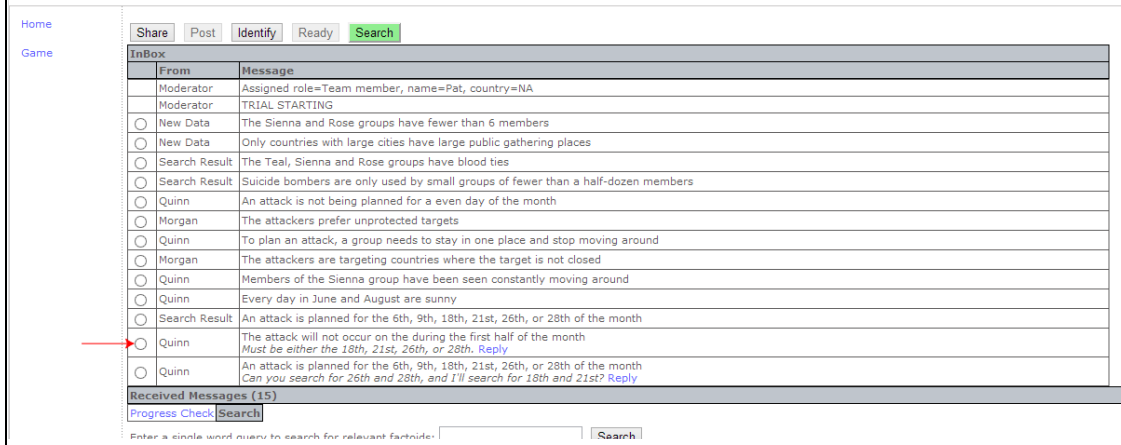
b) **Search**: Every 60 seconds, a subject is able to conduct a web search on any word the subject chooses in order to look for clues that may help to further the subject's theory of the who, what, where, and when of the attack. As shown in Figure A4, when the Search button is green, the subject may: (1) click on the Search button at the top of the page (which will be green 60 seconds after the past search or, conversely, will countdown in red the seconds remaining until 60 seconds has passed); (2) enter a term at the bottom of the screen; and (3) click on the adjacent Search button to execute the search. If a clue is found, it will report it out on the bottom of the screen (4) and add it to the subject's inbox (5). Search terms are full words—partial words will not work.

Figure A4: Search
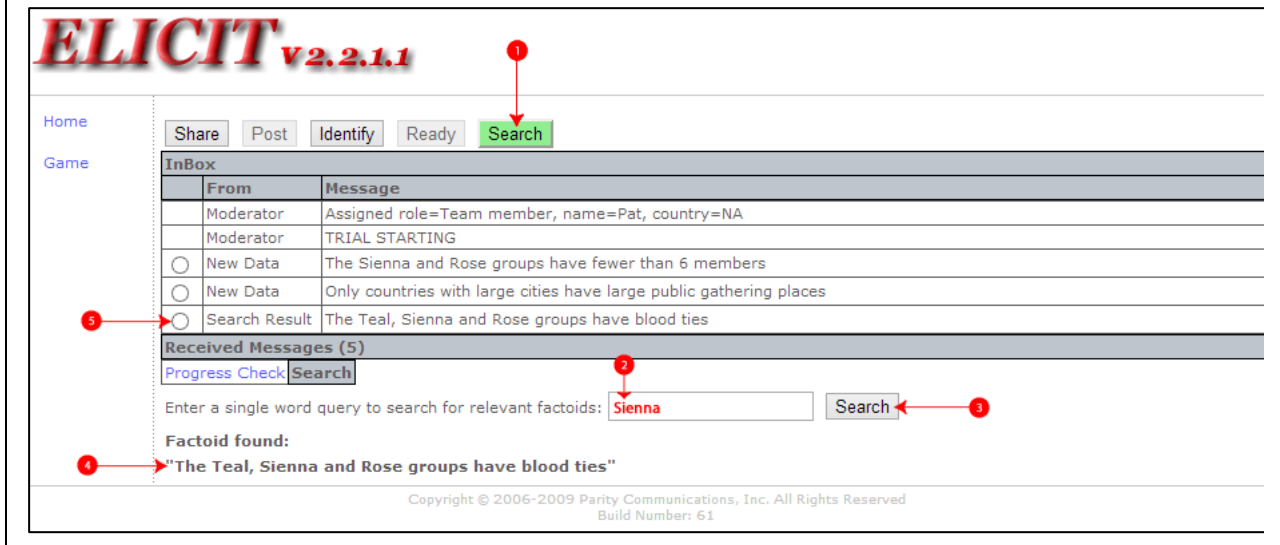
If the search term is not found in any of the remaining clues, the search will be unsuccessful, and the tool will return the message "No data associated with this key word" and the subject will be permitted to search again.

(c) **Identify / Progress Check**: A subject is paid based on how long the subject has the correct answer "identified" for any of the attack details submitted—either who, what, where, when in terms of month, when in terms of date, when in terms of hour, or when in terms of "Day" or "Night". A subject receives a specified rate per minute (15 cents per minute per sub-problem, or 3.75 cents per minute for each component of "when", for a maximum of 60 cents (a penny per second) for each minute) for each correct answer while it is submitted as that subjects first choice; there is no penalty for incorrect answers. A subject may submit and revise answers at any time, as many times as the subject wishes, by doing the following: (1) click on the "Identify" button at the top of the inbox; (2) enter (via direct entry or pull down) all of the answers the subject wishes to enter in the first choice line, and optionally, the confidence level (percent) for each answer along with any second-choice answers and confidence levels (which do not impact payment and are not shared with others); (3) click on "OK" to submit. (See Figure A5.)

Figure A5: Identify/Progress Check

After clicking "OK," those answers will be accessible to that subject, and the subjects to which that subject is connected, by viewing the "Progress Check" box at the bottom of the screen, as shown in Figure A6.

Figure A6: Example of Progress Check After Identify

ELICIT V2.2.1.1

Home
Game

| | Share | Post | Identify | Ready | Search |

InBox

	From	Message
	Moderator	Assigned role=Team member, name=Morgan, country=NA
	Moderator	TRIAL STARTING
●	New Data	The attackers prefer unprotected targets
○	New Data	The attackers are targeting countries where the target is not closed
○	Quinn	An attack is not being planned for a even day of the month
○	Dale	The Teal, Sienna and Rose groups have blood ties
○	Dale	Suspicious activity has been reported in stadiums *stadiums could be involved* Reply
○	Dale	The target is either a secular school, church, or army base *Nope -- scratch that* Reply
○	Pat	The attack will not occur on the during the first half of the month
○	Pat	An attack is planned for the 6th, 9th, 18th, 21st, 26th, or 28th of the month *Must be either the 18th, 21st, 26th, or 28th.* Reply
○	Pat	The Teal, Sienna and Rose groups have blood ties
○	Pat	Suicide bombers are only used by small groups of fewer than a half-dozen members
○	Robin	All groups have the capacity to operate in Muland, Xiland, Omicronland, Piland and Sigmaland
○	Quinn	An attack is not being planned for a even day of the month

Received Messages (14)

Progress Check

	Who	What	Where	Month	Day	Hour	AM/PM
Dale	Sienna - 5%	schools - 50%	Muland - 50%	Don't Know - 100%	Don't Know - 100%	6 - 50%	day - 50%
Morgan	Silver - 80%	school - 100%	Omicronland - 40%	June - 60%	21 - 100%	9 - 50%	day - 80%
Pat	Sienna - 70%	School - 80%	Piland - 100%	August - 30%	18 - 70%	6 - 50%	night - 50%
Quinn	Rose - 70%	Stadium - 90%	Omicronland - 70%	June - 100%	26 - 60%	8 - 30%	day - 50%
Robin	Sienna - 100%	church - 50%	Omicronland - 25%	August - 50%	7 - 1%	Don't Know - 100%	day - 100%

Answers (5)

While a subject may elect to identify answers at any time, the system will also prompt the subject to consider identifying answers at specified intervals throughout the session by automatically bringing up the pop-up identify screen shown in Figure A5.