# Shared language in the team network-performance association: Reconciling conflicting views of the network centralization effect on team performance

**Ray E Reagans** and **Hagay Volvovsky**
MIT Sloan School of Management, Cambridge, MA, USA

**Ronald S Burt**
Bocconi University, Milan, Italy; University of Chicago, Chicago, IL, USA

## Abstract
We reconcile two conflicting views of the network centralization effect on team performance. In one view, a centralized network is problematic because it limits knowledge transfer, making it harder for team members to discover productive combinations of their know-how and expertise. In the alternative view, the limits on knowledge transfer encourage search and experimentation, leading to the discovery of more valuable ideas. We maintain the two sides are not opposed but reflect two distinct ways centralization can affect a team's shared problem-solving framework. The shared framework in our research is a shared language. We contend that team network centralization affects both how quickly a shared language emerges and the performance implications of the shared language that develops. We analyze the performance of 77 teams working to identify abstract symbols for 15 trials. Teams work under network conditions that vary with respect to centralization. Results indicate that centralized teams take longer to develop a shared language, but centralized teams also create a shared language that is more beneficial for performance. The findings also indicate that the highest performing teams are assigned to networks that combine elements of a centralized and a decentralized network.

## Keywords
Teams, network, centralization, shared language, problem-solving, performance, learning

**Corresponding author:**
Ray E Reagans, MIT Sloan School of Management, E62-382, 100 Main St, Cambridge, MA 02142, USA.
Email: rreagans@mit.edu

## Significance Statement

In the context of unfamiliar and complex tasks, the challenge of how to organize teams effectively often leads to a choice between decentralized and centralized networks. Decentralized networks, such as fully connected or random networks, foster coordination, while centralized networks like hub and spoke or wheel networks support enhanced learning. However, when confronted with unfamiliar tasks, it remains uncertain whether coordination or learning holds greater importance, thus making it unclear which network structure would best facilitate superior team performance.

To resolve this debate, we focus on a team's shared language. Scholars on both sides of the debate agree that a team's shared language is critical for success. For those studying learning, a shared language embodies a team's code, encapsulating the collective knowledge of current best practices. Conversely, for those emphasizing coordination, it constitutes an integral component of a team's transactive memory system, encompassing awareness of each team member's expertise and actions.

Through our argument and presented evidence, we establish that network centralization has a dual impact on a team's shared language development and its subsequent performance implications. These two network effects combine to shape overall team performance and determine the superiority of either a decentralized or centralized network within a specific context. The findings advance our understanding of how network centralization contributes to team performance.

## Introduction

Scholars from several fields and disciplines have documented the effect network centralization can have on the performance of groups and teams (Balkundi and Harrison, 2006; Mathieu et al., 2008: 440–441; Park et al., 2020: 1010–1012). Network centralization is thought to be especially important when a team has been given an unfamiliar and complex assignment (Brass et al., 2004). The relationships in a centralized network (e.g., a wheel or hub and spoke network) are concentrated on one or a small number of individuals, while the relationships in a decentralized network (e.g., a fully connected or random network) are distributed more equally (Shaw, 1964: 122–123; Freeman 1977: 39). Scholars agree that network centralization is important when a team is tasked with a complex and unfamiliar assignment. They disagree, however, regarding whether a centralized or decentralized network is more likely to result in superior performance.

Scholars who expect for decentralized teams to be more productive emphasize the importance of coordination (Shaw 1964: 122–124; Cummings and Cross, 2003; Katz et al., 2004: 318; Balkundi and Harrison, 2006: 60; Mukherjee, 2016; Mora-Cantallops and Sicilia, 2019). The limits on direct communication in a centralized network make it difficult for team members to utilize their know-how and expertise (Huang and Cummings, 2011; Sherf et al., 2018). The relatively large number of direct and indirect communication channels in a decentralized network provides team members with an opportunity to quickly share their successes and failures and discover productive combinations of their individual knowledge and expertise (Huang and Cummings, 2011; Argote et al., 2018; Sherf et al., 2018).

Scholars who expect for centralized teams to be more productive emphasize the significance of learning (Lazer and Friedman, 2007; Fang et al., 2010; Csaszar and Siggelkow, 2010; Schilling and Fang, 2014). The research is influenced by March's (1991) discussion of exploitation (e.g., selection, refinement, and execution) and exploration (e.g., search, variation, discovery, and innovation) for organizational learning. The research illustrates how network structure determines if members of a collective will adopt a problem-solving strategy characterized by exploitation or exploration. The results illustrate a network-task contingency. If the unfamiliar assignment is basic, superior performance is more likely in a decentralized network (Lazer and Friedman, 2007: 682). The dense web of interconnections in a decentralized network facilitates knowledge transfer and social learning. When individuals are performing the same task, knowledge transfer represents exploitation (Argote and Ingram, 2000: 150). Individuals learning in a decentralized network can learn from each other and quickly discover and make incremental improvements to a promising idea.

If the assignment is complex, exploitation can undermine performance. The rapid exchange of ideas can increase short-term performance but could also limit longer-term success (Lazar and Friedman, 2007: 678–680). Discovering superior solutions to a complex task requires extensive search and experimentation, referred to as slow learning or exploration. A centralized network promotes exploration. The limited opportunities for communication and knowledge transfer among individuals in a centralized network restrict the individual's ability to learn from each other. Team members who are disconnected from each other can

focus on improving a distinct idea or solution, leading to system-wide diversity. Maintaining system-wide diversity while learning increases the likelihood that members of a collective will either discover or create a solution that generates higher performance outcomes (Hong and Page, 2004; Page, 2007; Lazer and Friedman, 2007; Maroulis et al., 2020).[1]

We attempt to reconcile the conflicting predictions for the network centralization effect on team performance by considering how network centralization affects the *emergence* and the *value* of a team's or group's shared problem-solving framework. While the two sides expect for different network structures to produce better outcomes, both sides agree that it is important for teams to develop a shared problem-solving framework. For scholars who study teams, the shared framework could be a transactive memory system (Lewis, 2003; Ren and Argote, 2011). For scholars who study learning, the framework could be a team code (Koçak and Warglien, 2020). Both sides agree that a shared framework facilitates coordination and learning. A shared framework is a stock of knowledge, or an institutional memory of the performance outcomes produced by past choices and decisions and provides a foundation for making better choices and decisions (Denrell et al., 2004). Network centralization can affect how quickly team members develop a shared framework and network centralization can also affect the performance implications of the framework team members develop. The rate a shared framework emerges is an example of the dynamics emphasized by researchers who highlight coordination, while the performance implications of the shared framework is an example of the outcomes emphasized by researchers who privilege learning.

The predictions described above are examined with a symbol identification task. The symbols are tangrams. Tangrams are abstract symbols that originated in China many hundreds of years ago. Several thousand tangrams can be constructed from seven generative shapes and each individual tangram can be described with many words and phrases. Team members are asked to identify a shared symbol from a set of symbols. Identifying abstract symbols is an unfamiliar and complex task. Since a symbol can be described with many words and descriptions, a team must develop a shared language to distinguish the symbols and select the symbol they have in common (Clark and Wilkes-Gibbs, 1986:11; Selten and Warglien, 2007). A shared language is a team code (Koçak and Warglien, 2020) and is also a part of a transactive memory system (Moreland et al., 1996).

The predictions we examine are illustrated in Figure 1. Network centralization is expected to have a negative effect on how quickly team members can develop a shared language. The limits on communication in a centralized network should make it difficult for team members to agree on a name or description for a symbol. The limits on communication should result in the more careful consideration of alternative names and the eventual selection of a name that helps team members identify the shared symbol more frequently. Teams that can more frequently identify the shared symbol are more accurate. To preview the results, teams are randomly assigned to either a centralized or decentralized network. The findings provide support for Figure 1. A shared language has a positive effect on team accuracy. Centralization has a negative effect on the rate a shared language emerges. But centralization also increases the magnitude of the positive effect a shared language has on team accuracy. Finally, teams are also assigned to networks that combine features of a centralized and decentralized network. Teams assigned to a network with both features were able to develop a shared language as fast as a decentralized team and the magnitude of the shared language effect on team accuracy was as large as the effect observed for centralized teams. The teams assigned to a network with both features were the highest performing teams.

Our research makes two contributions. We distinguish the network centralization effect on the rate a team develops a shard language from the network centralization effect on the value of the shared language a team develops. While previous research has examined the network centralization effect on the emergence of a shared language, the network centralization effect on the value of the shared language created by a team has not been considered. We introduce and document the network centralization effect on the value of a shared language developed by teams and demonstrate how the effect can help reconcile the conflict between scholars who expect higher performance in a centralized network and those who anticipate superior performance in a decentralized network.

## Centralization and team performance

In this section, we provide a detailed discussion of the links presented in Figure 1. Based on the findings reported by researchers who study coordination and those who study learning, we do not predict a main effect for centralization on performance. The centralization effect is assumed to not equal zero. The path from the shared language variable to team performance highlights the significance of a group or team developing a shared problem-solving framework. The shared framework represents a stock of knowledge that includes beliefs describing which ideas and solutions are worth pursuing. It can take the form of an organizational or team code, describing the "current" best practice. The shared framework can be a transactive memory system, incorporating beliefs that describe what team members know and do. The shared framework plays a critical role in facilitating coordination and learning within the team.
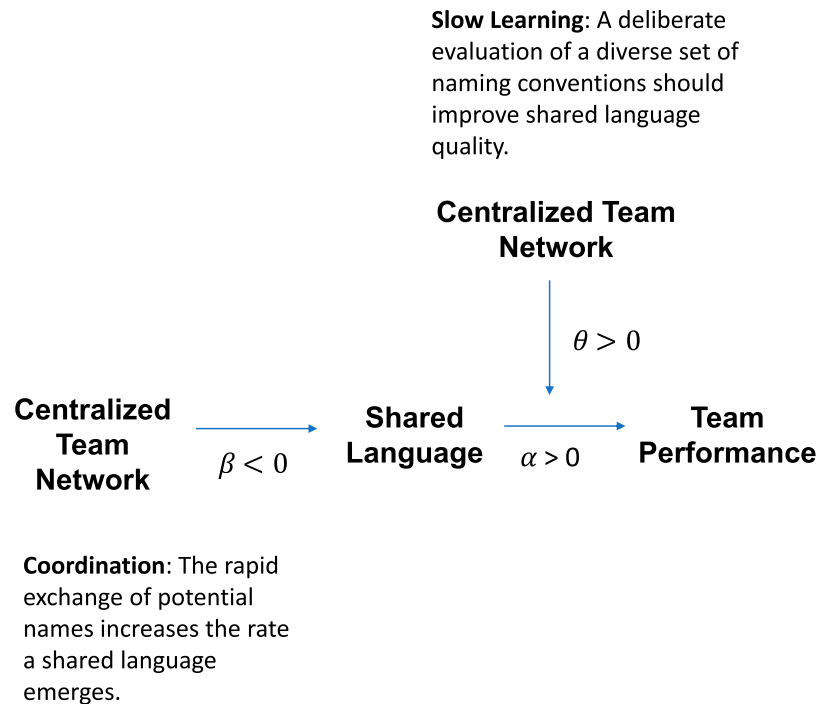
**Slow Learning**: A deliberate evaluation of a diverse set of naming conventions should improve shared language quality.

**Centralized Team Network**

$\theta > 0$

**Centralized Team Network**  →  **Shared Language**  →  **Team Performance**

$\beta < 0$                            $\alpha > 0$

**Coordination**: The rapid exchange of potential names increases the rate a shared language emerges.

**Figure 1.** Shared language in the team network-performance association.

The shared framework studied in our research is a shared language. As team members collaborate and accumulate experience working together, they often develop their own language and shorthand terms to describe elements of their work (Weber and Camerer, 2003). This shared language serves as a valuable tool for communication (Krauss and Fussell, 1991), enabling team members to assess different solutions' merits and effectively coordinate their efforts once potential solutions are identified. A team's shared language is an integral part of a transactive memory system (Moreland et al., 1996), known to enhance team performance (Weber and Camerer, 2003; Reagans et al., 2016). The shared language's positive impact on performance lies in its ability to facilitate coordination and learning within the team.

Based on previous research, we expect for a team's shared language to improve with experience, leading us to expect the emergence of a shared language to follow a team experience curve. The network centralization effect we discuss in the next section should be understood as an adjustment to the baseline team experience effect. The network effect interacts with and modifies the impact of team experience on the development of a shared language within the team.

### Centralization and the emergence of a shared language

The path from network centralization to a team's shared language can be traced back to the lab experiments conducted at MIT in the 1950s by Bavelas (1948, 1950), Smith (1950), and Leavitt (1949, 1951). These experiments examined the relationship between a team's communication network and its performance. Shaw (1964) provides an extensive review of the findings from 36 initial MIT experiments, which encompassed a range of communication networks, from those organized around a single individual to fully connected networks. The teams were engaged in various tasks, including basic symbol identification tasks (e.g., stars, squares, and circles) and more complex decision-making tasks (e.g., math problems, sentence completion problems, and discussion problems). Consistent differences were observed between centralized and decentralized networks. The research findings indicate that for basic problem-solving tasks, team performance was found to be higher in a centralized network. However, when teams tackled more complex problems, the pattern was reversed, and team performance was higher in decentralized networks (Shaw 1964:122–124).

In addition to establishing the network-task contingency emphasized by researchers who study teams, these early experiments also shed light on the relationship between coordination and learning. While coordination and learning are conceptually distinct, successful coordination becomes dependent on learning when the task at hand is unfamiliar and complex (Knudsen and Srikanth, 2014). An example of this is seen in the teams from one of the original MIT experiments that were tasked with identifying multicolor or "noisy" marbles (Christie et al., 1952). To accurately identify these "noisy" marbles, team members had to

develop a shared language. The teams had to learn how to coordinate their efforts. The development of a shared language increased team accuracy, and it was observed that teams operating within a decentralized network were more likely to develop a shared language.

Given prior research, we expect for our decentralized teams to develop a shared language at a faster rate. It is useful however, to describe the process we expect for decentralized teams to follow that will allow them to develop a shared language at a faster rate. The teams we study were asked to identify abstract symbols. The problems they are solving is identifying abstract symbols. The solutions to the problems are naming conventions (Centola and Baronchelli, 2015; Guilbeault et al., 2021). To identify the symbols, team members experiment with alternative naming conventions. The value of a naming convention lies in its ability to accurately identify a specific symbol. A shared language is simply a collection of naming conventions.

The set of potential naming conventions and their value is unknown to a team initially. The teams discover the potential set of naming conventions and the value of those conventions through trial-and-error experimentation. During a trial, team members select a set of naming conventions and use that set to identify the shared symbol. If their attempt is successful, the set of naming conventions is more likely to be retained. If their attempt is not successful, the team is expected to consider an alternative set of naming conventions (Baronchelli, 2016: 6). The teams can be expected to "brainstorm" as they puzzle their way through alternative naming conventions and modify those alternatives as they receive feedback on their choices and decisions.

Teams that work in a decentralized network can be expected to develop a shared language at a faster rate. The dense web of interconnections provides team members with an opportunity to learn potential naming conventions from each other. A potential naming convention is likely to be shared across multiple communication channels increasing the odds that it will be adopted by the team.[2] The rapid exchange of potential naming conventions increases the likelihood of adoption and therefore the rate at which a team can develop a shared language. The previous discussion along with the results from the experiment by Christie and his colleagues provides the foundation for our first prediction.

**Hypothesis 1.** Team network centralization reduces the positive effect team experience has on the emergence of a shared language.

## Centralization and the emergence of a more valuable shared language

There is a downside to the greater potential for social learning facilitated by the dense web of interconnections in a decentralized network. Indeed, the coefficients in Figure 1 indicate a tradeoff between the rate a shared language emerges and the value of the shared language that develops. Rapid convergence during a problem-solving process can reduce the quality of the ideas that are produced and discovered. For example, the research on brainstorming indicates that when compared to individuals who brainstorm alone, individuals who brainstorm together produce relatively inferior ideas. Individuals brainstorming together learn from each other and adjust their ideas and innovations in the direction of the ideas and innovations shared by their colleagues. The narrowing of the ideas under consideration reduces the overall quality the ideas that are produced (Rajaram and Pereira-Pasarin, 2010; Barber et al., 2015).

We expect for teams that adopt a problem-solving process that fosters diversity in potential naming conventions to produce conventions that are more valuable and beneficial for team performance. A diverse problem-solving process is more likely in a centralized network. The disconnects in a centralized network restrict communication and knowledge transfer, limiting the opportunities for team members to discover potential naming conventions from each other. The limitations on social learning should encourage individual search and experimentation, leading to greater diversity in the naming conventions under consideration. With more diversity in the naming conventions under consideration, the team is in a better position to select and retain more valuable naming conventions.

Figure 2 contains the tangrams we will use in our study. The six symbols are taken from a study of language-based coordination (Clark and Wilkes-Gibbs, 1986: 11). Beneath each symbol is a sample of descriptive names subjects used during the final five trials of our experiment. The three symbols on the bottom row are frequently described as a "bunny," which is not surprising. The word "bunny" or "rabbit" can be used to describe almost any abstract symbol (Guilbeault et al., 2021: 2; Burt and Reagans, 2022: 17). If a member of a decentralized team recommends the name bunny for the fifth symbol, the name is likely to be adopted by the team. The name bunny, however, is ambiguous and so is less valuable as a naming convention. It is possible a team member assumes the focal team member is talking about the fourth or sixth symbol.

Even if the team members discover that "bunny" is an ambiguous name and search for alternatives, their subsequent choices are more likely to be influenced by their initial choices. They may decide to use "bunny" for the fourth symbol, "bunny rabbit" for the fifth symbol, and "rabbit" for the sixth symbol. The general point is that a decentralized network limits exploration, and by limiting exploration reduces the value of the naming conventions the team selects and therefore the overall value of the shared language it develops.

A centralized network allows team members to search for and experiment with a broader set of alternatives.
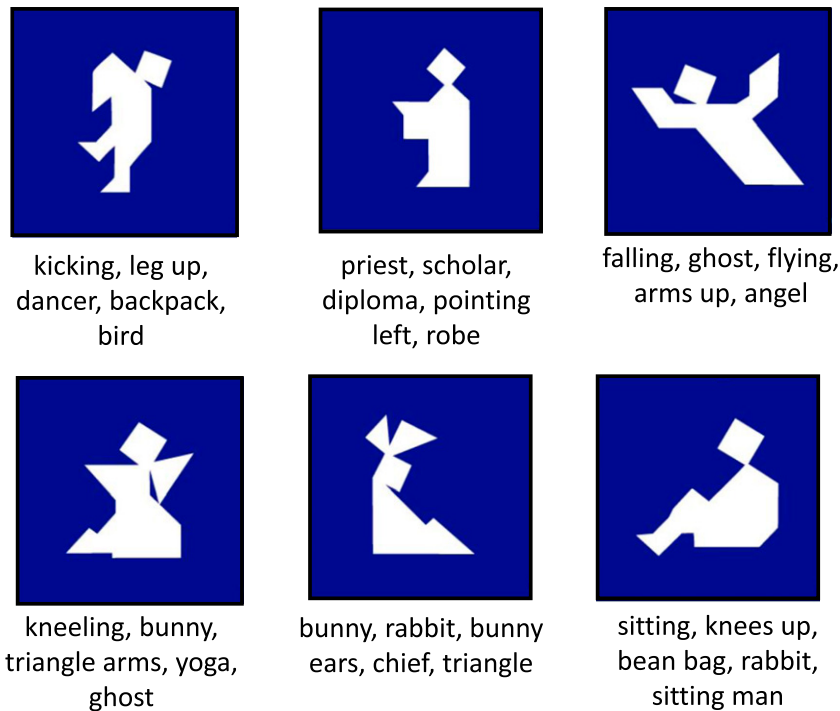
| | | |
|---|---|---|
| kicking, leg up, dancer, backpack, bird | priest, scholar, diploma, pointing left, robe | falling, ghost, flying, arms up, angel |
| kneeling, bunny, triangle arms, yoga, ghost | bunny, rabbit, bunny ears, chief, triangle | sitting, knees up, bean bag, rabbit, sitting man |

**Figure 2.** Abstract symbols and descriptive names.

A wheel network is an example of a centralized network. The central individual in the wheel network has more power and controls the flow of knowledge and information (Burt, 2021). The power and influence of the central team member is one reason why we expect for centralized teams to generate more valuable shared languages. The peripheral members of a centralized network also play an important role. For example, if the central member decides to select "bunny" for the fourth symbol, the peripheral members can continue to refer to the fourth symbol as "chief," the "peace sign," or "bunny ears" until someone on the team suggests naming it "graduate," which is a more valuable naming convention because it is less likely to be used for other symbols. The central member can decide that "graduate" is an effective name and share it with the rest of the team.

There are two reasons why a centralized network can produce a more valuable shared language. A centralized network facilitates system-wide diversity. The disconnects between peripheral team members limit their opportunities to learn potential names from each other, resulting in a higher likelihood that peripheral members will maintain distinct naming conventions. A centralized network also provides a strong foundation for selection. The central team member can evaluate the alternatives introduced by the peripheral members and select and share promising naming conventions with the team. A centralized network enables the variation and selection that often leads to the discovery or creation of more valuable ideas and innovations. This line of reasoning forms the basis of our second prediction.

**Hypothesis 2.** Team network centralization increases the magnitude of the positive effect a shared language has on team performance.

## Network experimental conditions

We evaluate the predictions in four network structures. The networks are illustrated in Figure 3. The wheel is a centralized network, and the fully connected (FC) network is a decentralized network. A team working in a wheel network is a wheel team and a team working in a FC network is a FC team. We modify the wheel network to create two intermediate networks. The wheel network can be modified to include two central individuals instead of one (Rogge, 1953: 18). The presence of an additional central individual introduces more indirect communication channels between the remaining team members. More indirect communication channels could enhance the team's ability to reach consensus and develop a shared language. But given limits on team size, introducing an additional central individual would come at the expense of a distinct perspective on the team. The benefits of slow learning would be diminished but should not be completely offset.

The central individual in the wheel is commonly referred to as a broker. We call the intermediate networks two broker networks, and two versions of the two-broker network are considered. In one structure, the two central individuals are not connected. We call this network the disconnected brokers (DB) network and a team with a DB network is a DB team. In the other structure, the
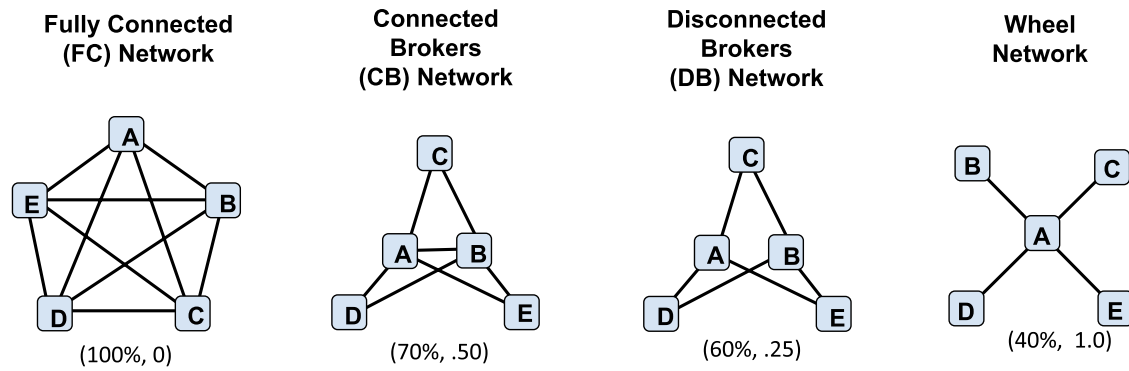
**Figure 3.** Four team networks.

two central individuals are connected. We call this network the connected brokers (CB) network and a team with the CB network is a CB team. The connection between the two brokers in the CB network is the critical difference between the CB and DB networks. The direct connection between the two brokers creates the potential for more effective coordination between the two brokers which should create an even greater capacity for developing a shared language on a CB team. However, we do not know how much experience working together the two brokers will need before they learn how to work together effectively.

We have focused our discussion on centralization, but the four networks also differ with respect to degree and network density. While centralization and density are distinct concepts, the two structural features are interrelated (Butts, 2006). In general, centralization declines as network density increases. The available research indicates that centralization is a complex function of network size (team size in our case) and network density (Anderson et al., 1999). Anderson and his colleagues indicate that centralization should be understood relative to network size and density. They maintain: "Unlike many interactions familiar to data analysts in the social sciences, these interactions are both fundamental and non-trivial: generally speaking, they can neither be removed through judicious experimental design, nor can they be accounted for simply by adding a covariate to a regression equation (pg.: 258)." Any explanation that focuses on centralization must also include network size and density as explanations. Network size in our study is fixed, but network density varies. While we have focused on centralization in our argument, it is important to remain mindful of network density, as we will see in the results sections.

## Experimental design

The experiment was conducted in laboratories at two Universities in New England. The subject pool contains students from the two schools, and neighboring schools, along with adults from the surrounding communities. Pretesting indicates that non-native English speakers find the symbol identification task difficult to complete with native speakers, and older subjects often have difficulty with features of the software used in the experiment, so participation is limited to native English speakers and people between the ages of 18 and 55.

The experiment is conducted at computers. The subjects arrive at a laboratory; are seated in separated computer cubicles; and then assigned to teams. The teams are randomly assigned to network conditions, and team members are randomly assigned to different positions on their team. The network conditions and positions are fixed in all trials of play.

Figure 4 is a subject's screen. In the upper-left part of the screen, there are five symbols in the "My Card" box which are the subject's symbols on the focal trial. The five symbols are a subset of the six symbols illustrated in Figure 2. There are six distinct combinations of the Figure 2 symbols taken five at a time. One symbol is shared in five of the possible six combinations, which allows each subject on a five-person team to receive a different set, with one common symbol across the five sets.

Team members communicate with dialog boxes. An individual communicates with a teammate by clicking on the teammate's number in the "My Network" dialog box at the top of the screen. The teammates listed in the dialog box indicate direct communication channels. The screen in Figure 4 is for player 2, who can communicate directly with all four teammates. As a subject communicates with teammates, a teammate-specific chat-box at the bottom of the screen accumulates the messages they exchange. For example, the messages to and from "Player four" are accumulated in the chat-box with the "Player four" header. The messages sent and received during a trial can be reviewed by moving the dialog-box slider up or down. The chat-box resets on each trial.

Subjects communicate about their tangrams until making a guess about the shared symbol. To submit his or her answer, the subject highlights one of the tangrams at the top of the screen and clicks the "Submit Answer" button below the tangrams. Dots at the top of the screen darken as teammates
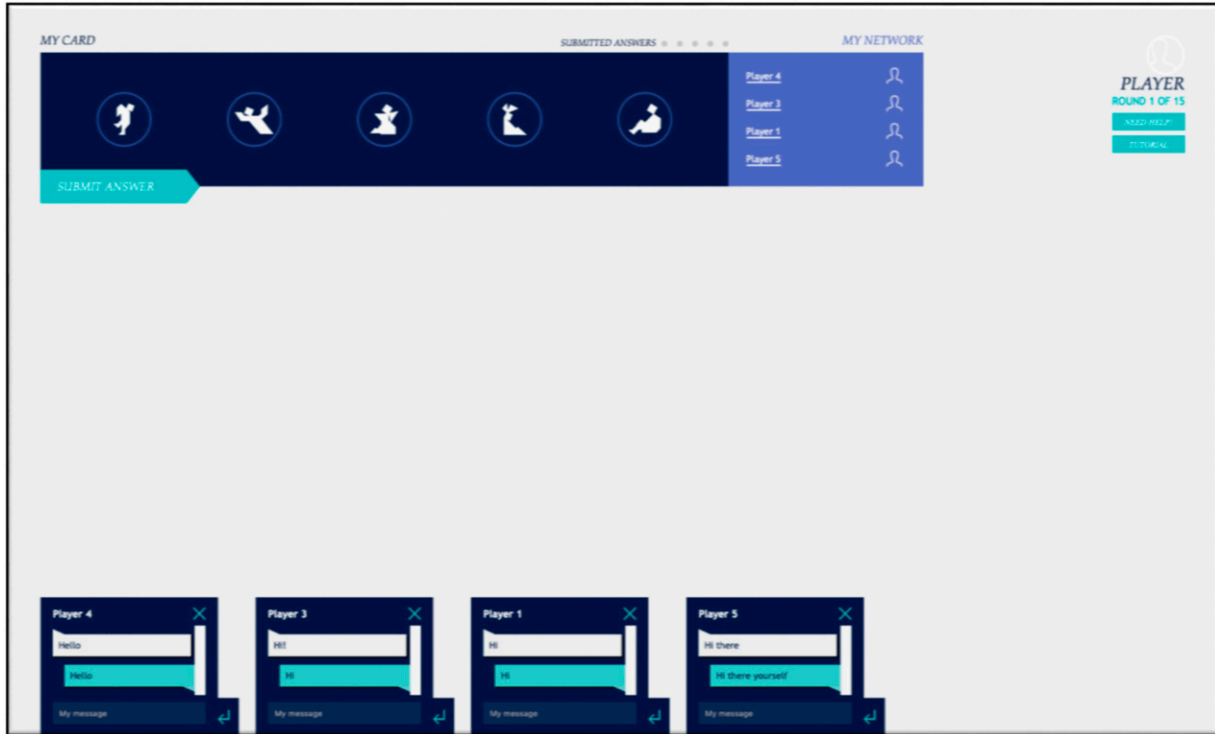
**Figure 4.** Example subject screen.

submit answers, so an individual knows how the team is progressing. Subjects do not see teammate answers, but the darkened dots allow them to see how many have submitted answers. When a subject submits an answer, the "Submit Answer" button on the screen turns into "Reconsider." After communicating with team members, an individual could decide that he or she selected the wrong shared symbol and would like to reconsider his or her answer. If "Reconsider" is clicked before all other teammates submit answers, the number of darkened dots decreases by one. A trial ends when all five people have submitted their answer.

Feedback is immediate and immediate feedback facilitates learning (Burgess, 1968: 327; Sutton and Barto, 1998; Selten and Warglien, 2007). If everyone correctly guesses the shared tangram, "Correct" shows on the screen. One or more incorrect guesses yields "Incorrect." After feedback is given, the screen clears, each subject receives a new card, and the next trial begins. The experiment lasted for 15 trials. Teams were given 75 min to finish all 15 trials. Data on 77 teams (i.e., 385 men and women) was collected.

### Shared language

To calculate the shared language variable, we use recent developments in natural language processing to define the meaning of the words team members use while communicating. The recent developments quantify the famous quotation by Firth (1957) which states: "You shall know a word by

the company it keeps." Meaning is derived from context. Using this basic idea, scholars have attempted to represent each word as a vector. The approach is called word to vector or word2vec (Mikolov et al., 2013; Bojonowski et al., 2017). With word2vec, a word's meaning is derived from the words that it occurs with more frequently. A detailed description of how words are represented as vectors is beyond the scope of the current discussion. We note, however, that two words will have similar vector representations, and therefore meaning, if they occur frequently with the same words (Levy and Goldberg, 2014: 4-5). We use fasttext to create a vector for every word in the message data. fasttext is an open-source library for text representation and classification (Bojonowski et al., 2017). We follow convention in excluding stopwords from our text analysis. Stopwords are words which occur more frequently in text and so have ambiguous meaning.

Each message is represented as a vector by taking the average of the word vectors it contains (Le and Mikolov, 2014).[3] Cosine similarity is used to measure the extent to which consecutive messages, $M_t$ and $M_{t+1}$, on a team trial contain words with similar meaning:

$$\text{shared language} = S_c(M_t, M_{t+1}) = \cos(\theta) = \frac{M_t * M_{t+1}}{\|M_t\| \|M_{t+1}\|}$$

(1)

The t indicator in the shared language equation is an indicator of message sequence during the trial and not an indicator of the trial. For each team trial, the shared language

scores are averaged to define the extent team members communicate with similar words and phrases during a trial.

## Estimation

To estimate the network effects in Figure 1, two estimation issues must be addressed. First, teams complete the task multiple times. There are 77 teams but since teams complete the task multiple times, there are 966 team trials. Clustering by teams violates the independence assumption in regression analysis. To adjust our estimates for clustering, we include in our models a random effect for each team and allow the team random effects to be correlated with predictors in our regression equations. A correlated random effect is a "fixed" effect (Mundlak, 1978; Wooldridge, 2010: 346-361). The fixed effect controls for unobserved differences between teams. Second, we expect for the network conditions to affect a team's shared language and for a team's shared language to affect team performance. A team's shared language is endogenous. To address both concerns, we estimate a structural equation model (SEM). The SEM framework allows us to include a random effect for each team and to allow the random effects to be correlated with our predictors. The SEM framework also allows for us to estimate the path from our network conditions to the shared language variable and from the shared language variable to team performance. Our indicator of team performance is team accuracy, which is a count of the number of teammates who correctly identify the shared symbol on the focal trial. The dependent variable in our performance equation is assumed to follow a binomial distribution with an upper limit of five, the size of our teams (Wooldridge, 2010: 739-740). The empirical results lead to the same substantive conclusions if our dependent variable is the proportion of correct responses.

The trial variable is included in the shared language and the team accuracy equations. A large body of research has documented the positive effect team experience can have on team processes and performance outcomes (Reagans et al., 2005; Huckman et al., 2009). Experience working together increases across trials and is expected to affect the emergence of a shared language but can also affect team accuracy. We also include several control variables. Preliminary analysis indicates some symbols are harder to identify, so we control for the shared symbol on a trial. 77 teams started the experiment, but 32 teams did not complete all 15 trials. Teams that do not finish all 15 trials collapse. Teams vary in terms of when they collapse. No teams collapse during the first four trials. Four percent collapsed between trials five and nine. The percent increases to 17 percent on trial 10 and declines to three percent during the final five trials. We include two control variables to capture this dynamic. A dummy variable is included for the tenth trial. We also include a team collapse variable that is

set equal to zero and changes to one if the focal trial is the terminal trial for a team that collapses. Finally, we control for where the data was collected. We include a binary variable set equal to one if the data was collected at the first laboratory and is set equal to zero if the data was collected at the second laboratory. Summary statistics for the variables in the analysis are in Table 1.

## Results

Figure 5 contains the distribution of the shared language variable across the four network conditions. The average level of the shared language variable is higher in the FC (0.831) and CB (0.845) networks than in the DB (0.818) and wheel (0.815) networks. Figure 6 contains the distribution of the team accuracy variable. Team accuracy is a count variable and indicates the number of team members who correctly identify the shared symbol on the focal trial. Team accuracy varies from zero to five. The tendency is for everyone or no-one on the team to correctly identify the shared symbol. The CB teams are the most accurate, while the FC teams are the least accurate. The FC and CB teams exhibit similar levels of the shared language variable, but the teams differ with respect to team accuracy.

The shared language and team accuracy variables are the dependent variables in our regression equations. The results from the regression analysis are presented in Table 2. There are two sets of results in Table 2. Model 1 does not include control variables. Model 2 does. The coefficients for our focal variables change very little with or without the controls. We focus on the coefficients from model 2. The regression results allow us to describe how much any observed differences in team accuracy, for example, between the FC and CB teams, can be attributed to differences in the rate teams develop a shared language and the magnitude of the shared language effect on team accuracy. The wheel network is the excluded network in the regression equations. The regression coefficients for the interactions between the network indicators and the trial variable in the shared language equation are relative to the wheel network. The regression coefficients indicate a shared language emerges at a faster rate in the FC, CB, and DB networks than in the wheel network. The regression coefficients for the FC, CB, and DB networks are positive and significant. To compare the regression coefficients for the FC, CB, and DB networks, we use the MARGINS command in STATA.[4] The tests indicate the coefficient for the FC network is not significantly different from the coefficient for the CB network or the coefficient for the DB network. The coefficient for the CB network, however, is larger than the coefficient for the DB network.

The interactions between the network conditions and the trial variable are plotted in Figure 7. The solid line in Figure 7 is the predicted level of the shared language

**Table 1.** Summary statistics.

| | Mean | Std. Dev | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) | (13) | (14) | (15) | (16) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| (1) Team accuracy | 3.831 | 1.947 | 1.000 | | | | | | | | | | | | | | | |
| (2) Shared language | .827 | .050 | 0.257 | 1.000 | | | | | | | | | | | | | | |
| (3) Symbol1 | .143 | .350 | 0.041 | 0.020 | 1.000 | | | | | | | | | | | | | |
| (4) Symbol2 | .173 | .378 | −0.007 | −0.010 | −0.187 | 1.000 | | | | | | | | | | | | |
| (5) Symbol3 | .184 | .388 | 0.018 | 0.017 | −0.194 | −0.217 | 1.000 | | | | | | | | | | | |
| (6) Symbol4 | .145 | .352 | −0.143 | −0.038 | −0.168 | −0.188 | −0.196 | 1.000 | | | | | | | | | | |
| (7) Symbol5 | .186 | .390 | −0.006 | 0.013 | −0.195 | −0.219 | −0.227 | −0.197 | 1.000 | | | | | | | | | |
| (8) Symbol6 | .169 | .375 | 0.090 | −0.003 | −0.184 | −0.206 | −0.214 | −0.185 | −0.216 | 1.000 | | | | | | | | |
| (9) Laboratory one | .326 | .469 | −0.061 | −0.033 | 0.044 | −0.014 | −0.006 | 0.027 | −0.015 | −0.030 | 1.000 | | | | | | | |
| (10) CB network | .222 | .415 | 0.068 | 0.198 | −0.018 | 0.007 | −0.016 | −0.036 | 0.033 | 0.026 | −0.291 | 1.000 | | | | | | |
| (11) DB network | .282 | .450 | 0.013 | −0.111 | 0.014 | 0.085 | −0.084 | −0.003 | −0.004 | −0.006 | −0.013 | −0.334 | 1.000 | | | | | |
| (12) FC network | .259 | .438 | −0.092 | 0.052 | 0.056 | −0.108 | 0.067 | −0.008 | 0.009 | −0.014 | 0.391 | −0.315 | −0.370 | 1.000 | | | | |
| (13) Wheel network | .238 | .426 | 0.015 | −0.130 | −0.055 | 0.014 | 0.035 | 0.046 | −0.037 | −0.005 | −0.104 | −0.298 | −0.350 | −0.330 | 1.000 | | | |
| (14) Log trial | 1.741 | .767 | 0.317 | 0.382 | 0.017 | 0.003 | −0.018 | 0.004 | −0.016 | 0.013 | 0.000 | 0.005 | −0.015 | −0.024 | 0.036 | 1.000 | | |
| (15) Tenth trial | .065 | .247 | 0.066 | 0.091 | −0.012 | −0.032 | 0.026 | 0.022 | −0.029 | 0.027 | 0.013 | 0.000 | 0.002 | −0.003 | −0.000 | 0.193 | 1.000 | |
| (16) Team collapse | .033 | .179 | −0.180 | −0.046 | −0.009 | 0.007 | −0.028 | −0.010 | 0.030 | 0.009 | 0.007 | −0.015 | 0.026 | 0.023 | −0.036 | 0.103 | 0.209 | 1.000 |

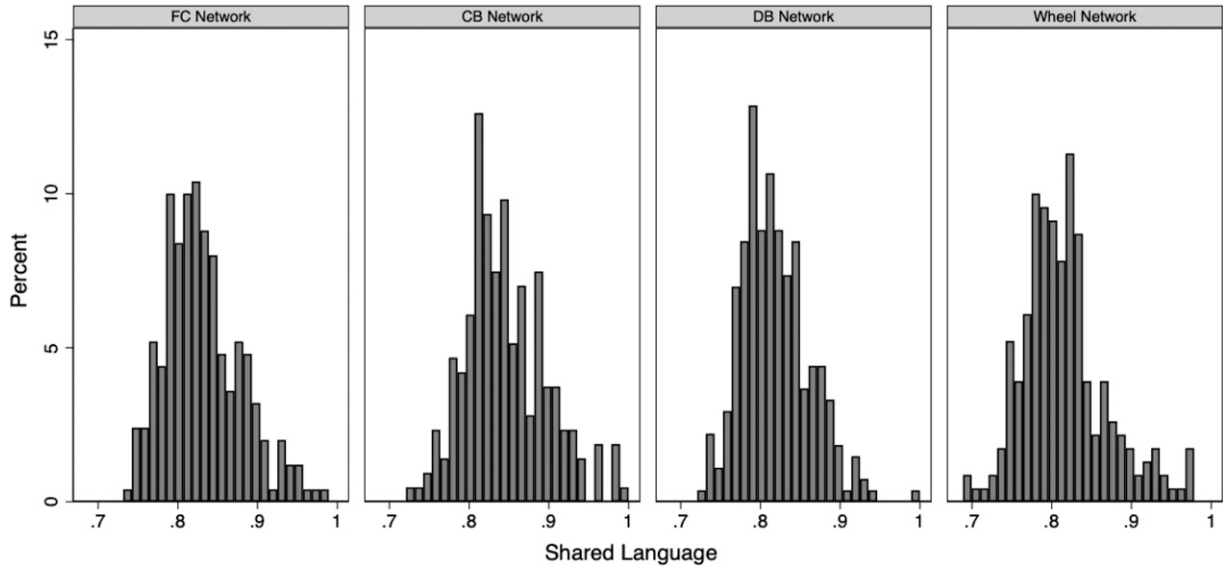Correlations with a magnitude greater than .066 are significant at the .05-level.

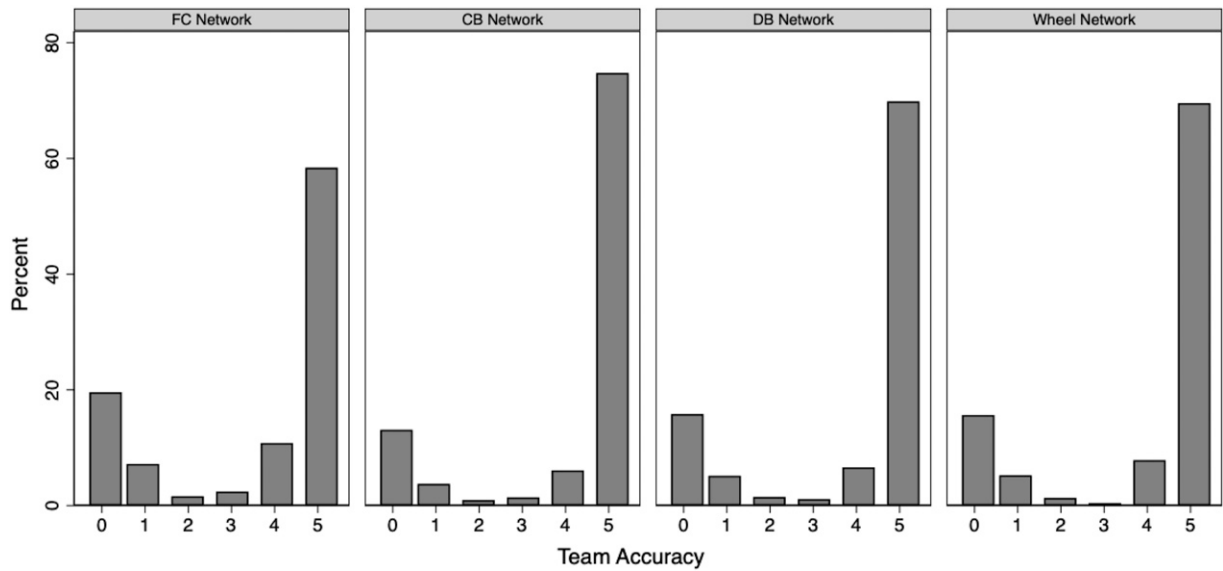**Figure 5.** Distribution of shared language by network structure.



**Figure 6.** Distribution of team accuracy by network structure.

variable for each trial. The shaded area is the 95% confidence interval for the predicted effect. As expected, the level of the shared language variable is lower in the wheel network than in the other networks. There appears to be some level differences between the FC, CB, and DB networks. We compare the levels of the shared language variable for the FC, CB, and DB networks trial-by-trial. Statistical tests indicate that after the 5th trial, the level of the shared language variable is significantly higher in the CB network than the DB network. After the 8th trial, the level of the shared language variable in the FC network is higher than

the level in the DB network, but the difference is only significant at the 10% level, which is only marginally significant. The level of the shared language variable is higher in the CB network than the FC network. The difference is close to but does not reach marginal significance.[5]

If we step away from the details, the observed outcomes provide support for the path from centralization to the shared language variable in Figure 1. The shared language variable emerges at the slowest rate in the most centralized network. But the results also illustrate why it is important to remain mindful of network density. The shared

**Table 2.** Network centralization, shared language, and team accuracy.

|  | Model 1 | Model 1 | Model 2 | Model 2 |
|---|---|---|---|---|
|  | Shared language | Team accuracy | Shared language | Team accuracy |
| Symbol1 |  |  | −0.001 (−0.30) | −0.303* (−2.20) |
| Symbol2 |  |  | −0.002 (−0.53) | −0.135 (−0.99) |
| Symbol3 |  |  | −0.011* (−2.42) | −1.015** (−7.47) |
| Symbol4 |  |  | 0.001 (0.29) | −0.149 (−1.09) |
| Symbol5 |  |  | −0.005 (−1.04) | 0.313* (2.11) |
| Laboratory one |  |  | 0.003 (0.40) | 0.178* (2.02) |
| CB network | −0.005 (−0.43) | −1.827 (−0.76) | −0.007 (−0.53) | −0.467 (−0.19) |
| DB network | −0.014 (−1.23) | 3.529 (1.63) | −0.014 (−1.20) | 2.312 (1.02) |
| FC network | −0.008 (−0.64) | 6.923** (3.51) | −0.006 (−0.54) | 8.063** (3.83) |
| Log trial | 0.011** (3.41) | 0.693** (14.53) | 0.011** (3.47) | 0.809** (15.74) |
| CB X log trial | 0.020** (4.20) |  | 0.020** (4.29) |  |
| DB X log trial | 0.011* (2.52) |  | 0.011* (2.53) |  |
| CLIQUE X log trial | 0.015** (3.22) |  | 0.015** (3.28) |  |
| Shared language (SL) |  | 13.850** (7.37) |  | 13.170** (6.52) |
| CB X SL |  | 2.177 (0.73) |  | 0.462 (0.15) |
| DB X SL |  | −4.420 (−1.64) |  | −2.858 (−1.01) |
| Clique X SL |  | −9.164** (−3.76) |  | −10.490** (−4.03) |
| Tenth trial |  |  | 0.009 (1.70) | 0.841** (3.78) |
| Team collapse |  |  | −0.016* (−2.20) | −2.525** (−12.55) |
| Constant | 0.793** (91.98) | −11.050** (−7.36) | 0.794** (74.37) | −10.490** (−6.48) |

The wheel is the excluded and reference network. Z-scores in parentheses: *$p < 0.05$, **$p < 0.001$.
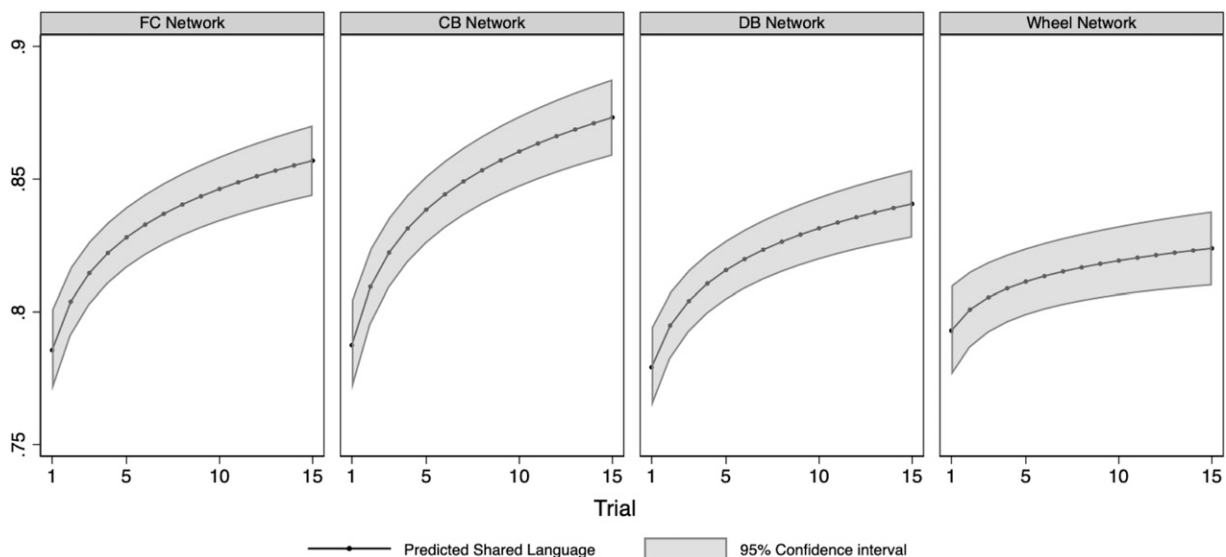


**Figure 7.** Emergence of shared language by network structure.

language variable emerges at faster rates as network density increases.

The regression coefficients for the interactions between the network indicators and the shared language variable describe how the magnitude of the shared language effect on team accuracy varies across the network conditions. The wheel network is the reference network. The coefficients for the CB and DB networks are not significant. The shared language effect in the CB and DB networks equals the shared language effect observed in the wheel network. The
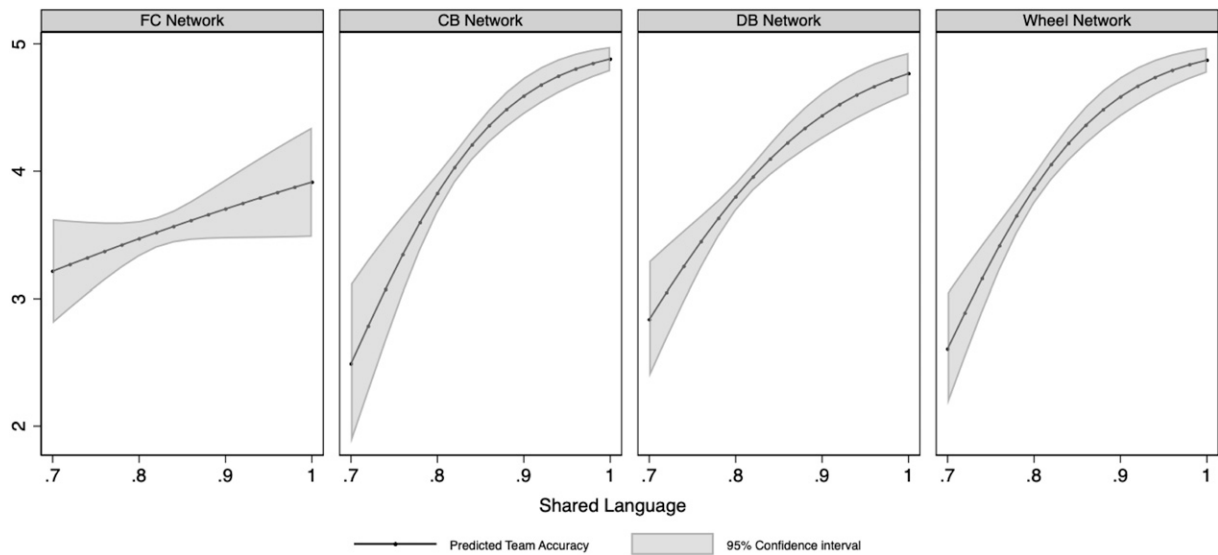
**Figure 8.** Shared language effect on team accuracy by network structure.

coefficient for the FC network is negative and significant. We use the MARGINS command to compare the regression coefficients for the shared language effect on team accuracy among the CB, DB, and FC networks.[6] The regression coefficients for the CB and DB networks are not significantly different from each other, but each one is larger than the regression coefficient for the FC network.

The interactions between the network conditions and the shared language variable are illustrated in Figure 8. The solid line in Figure 8 is predicted team accuracy for each level of the shared language variable. The shared area is the 95% confidence interval for the predicted effect. The shared language effect on team accuracy is significantly lower in the FC network. Team accuracy increases dramatically with the shared language variable in every network except the FC network. The 95% confidence intervals are informative. When the shared language variable is low, the width of the confidence intervals is relatively large, but declines as the shared language variable increases, except for the FC network. Some teams in the FC network with a high shared language are accurate and some teams with the same shared language level are relatively inaccurate. Teams in a FC network can develop a shared language, but the shared language they develop can sometimes have a modest effect on team accuracy.

It is again useful to step away from the details of the results. The outcomes provide support for the slow learning pathway in Figure 1. The magnitude of the shared language effect on team accuracy is larger for centralized teams. The results contrast outcomes in the teams with some centralization, the wheel, CB, and DB teams against the FC network that does not have any centralization.[7]

### The best team

The results indicate the network structures create distinct advantages. The FC and CB teams have an advantage in developing a shared language, while the wheel, CB, and DB teams have an advantage in creating a shared language that has a larger effect on team accuracy.[8] The two effects come together to define overall team accuracy. When we focus on overall accuracy, the CB teams are the best teams. The CB teams are more accurate than the wheel, DB, and FC teams. The differences are all statistically significant. The wheel teams are next. The wheel teams are more accurate than the DB and FC teams, and the DB teams are more accurate than the FC teams. The FC teams are the least accurate teams.

### Summary and discussion

Figure 1 illustrates how network centralization can affect how quickly a shared language emerges, and the performance implications of the language that develops. We find support for Figure 1. The figure is an attempt to reconcile conflicting predictions for the network centralization effect on team performance. Network centralization affects how quickly team members develop a shared language; an outcome consistent with the coordination benefits emphasized by scholars who expect for performance to be higher in a decentralized network. Network centralization also shapes the value of the shared language team members develop; an outcome consistent with the learning benefits highlighted by scholars who expect for performance to be higher in a centralized network.

While we find support for Figure 1, it is important to emphasize the empirical ambiguity the framework conveys.

For example, we can expect for a FC team to have an advantage in developing a shared framework and for a wheel team to have the edge in terms of the performance implications of the shared framework a team develops, but we cannot predict if overall performance will be higher in a FC network or a wheel network. Overall performance in either network will depend on the relative magnitudes of the coefficients in Figure 1, which are unknown a priori. There is no way to know which pathway will carry more weight in a specific setting, and therefore if a centralized or decentralized network will produce the best performance outcomes. More research is needed. With more research, we might discover that one pathway is always more consequential. But we could also discover that network centralization simply has mixed implications for team performance. The overall centralization effect could vary from positive to negative across research contexts. Figure 1 is an important first step. The specific results, however, are far from definitive.

While more recent is needed, our research findings are consistent with existing research. For example, the findings from our experiment are consistent with an emerging body of research emphasizing how sharing a language can affect individual performance (Srivastava et al., 2018), including the potential for an individual's network to shape the shared language effect on his or her attainment (Goldberg et al., 2016). The empirical results from previous research indicate that the shared language effect on individual performance varies with the network structure maintained by the focal individual. If the individual occupies a central position in his or her network, sharing a language improves performance. When network members spoke the same language, the central member was more likely to receive a favorable performance review and was less likely to involuntarily leave the organization. A shared language is more beneficial when the focal individual occupies a central position in his or her communication network.

Our results are also consistent with research emphasizing the superiority of network structures that combine features of a decentralized and a centralized network (Burt, 2000: 392–398; Reagans and Zuckerman, 2001; Reagans et al., 2004; Burt, 2005; Mehra et al., 2006; Lazer and Friedman, 2007; Fang et al., 2010). For example, research on teams in the field indicates that successful teams have an internal network that resembles a FC network, while the relationships team members have with colleagues from outside the team resemble a wheel, with the team as the hub in a hub and spoke network. It might appear that our results are inconsistent with these frameworks. But they are not. The CB teams in our experiment are the best teams. The CB teams allow for the benefits of centralization and decentralization to be realized. The CB teams, like successful teams in general,

strike a balance between creative problem-solving and a capacity for collective action. We are not suggesting the frameworks are the same but simply noting their family resemblance.

Our results also have important implications for scholars who study networks and knowledge transfer. Individuals who successfully bridge disconnects in network structure must synthesize complex knowledge and information (Burt, 2021). Successful knowledge transfer is the precursor to synthesis and individuals who bridge disconnects in network structure often acquire a capacity for knowledge transfer, which is especially critical when the knowledge being shared is complex (Reagans and McEvily, 2003; Tortoriello et al., 2012). The CB teams in our experiment illustrate how multiple people, in our case two, interpret, synthesize, and give meaning to the knowledge and information they receive.

The CB network is a "dual network" (Ter Wal et al., 2020). A dual network is a boundary spanning network that can facilitate knowledge transfer between areas of expertise. Ter Wal and her colleagues describe how a dual network can allow a manager and a scientist to combine scientific know-how and business expertise to produce superior innovations (Ter Wal et al., 2020). The two brokers at the center of the CB network are connected to each other but they also have "Simmelian" ties (Krackhardt, 1999) or "wide bridges" with the remaining team members. Multiple communication channels into the same information and knowledge sources can be expected to increase knowledge transfer (Tortoriello and Krackhardt, 2010) and improve outcomes that require the synthesis of complex knowledge and information. Complex knowledge and information can be interpreted in multiple ways. The CB teams create higher value interpretations.

Our research illustrates the value of opening the black-box between a group's network and how well a group can be expected to perform when given an unfamiliar and complex assignment. We do not, however, want to overstate the importance of the specific results. The teams work in a laboratory setting, for a relatively short period of time with strangers. One could ask how much can we really learn from a symbol identification task? One can imagine parallel activities. Tasks that are both unfamiliar and complex are often ill-structured (Simon, 1973). When a task is ill-structured, defining the problem and identifying potential solutions requires experience. Problem definition and potential solutions are unknown a priori and must be discovered while problem-solving. New product development can be an ill-structured task. We would expect for performance with respect to ill-structured tasks to vary, in part, with the dynamics illustrated in Figure 1. Given our results, more examinations of Figure 1 across a variety of task contexts would seem worthwhile.

## ORCID iD

Ray E Reagans ⬤ https://orcid.org/0000-0001-8706-3124

## Notes

1. The dense web of interconnections in a decentralized network and the disconnects in a centralized network shape opportunities for knowledge transfer. Motivation is also important, and motivation increases with tie strength. Strong network connections foster trust and cooperation while reducing interpersonal conflict and competition (Uzzi, 1999). The discussed network features can also influence motivation. Relationships are stronger when embedded in a dense web of interconnections (Rivera et al., 2010). The interconnections in a decentralization network promote knowledge transfer by increasing the reputation costs associated with withholding valuable knowledge and information (Reagans and McEvily, 2003: 245–246; Tortoriello et al., 2012: 1032; Reagans et al., 2015). Decentralized networks facilitate rapid knowledge transfer by providing more opportunities and greater motivation to share.

2. The effect network connections have the emergence of a shared language is an example of social contagion (Burt, 1987; Centola, 2018). An abstract symbol is an ambiguous stimulus. Network connections allow individuals to share their interpretations and eventually converge on a common name or description. The correct name is a social convention. The name is correct because everyone agrees it is the appropriate response. Centola and his colleagues use symbol identification tasks to study network effects on the emergence of social conventions (Centola and Baronchelli, 2015; Guilbeault et al., 2021). Findings indicate a social convention is more likely to emerge in a decentralized network.

3. Any information in the sequence of word use is lost with this approach. Results to be presented are the same if we use a more advanced method which includes this additional information in our message vectors (Le and Mikolov, 2014).

4. For example, we calculate the difference between the regression coefficients for the FC and CB teams ($\beta_{FC} - \beta_{CB}$). The difference, along with the standard errors of the regression coefficients are used to calculate a z-statistic. We report differences between coefficients that are significant at the 0.05-level.

5. We have focused on team members using similar words and phrases. The types of words teams decide to use are informative. We cluster analyzed our word vectors to identify words with similar meanings (Fraley and Raftery, 2002). The results indicate our teams sometimes describe the symbols using shapes (e.g., rhombus, trapezoid, and triangle), perceived actions (e.g., walking, sitting, and looking), or names or labels (e.g., man, priest, and rabbit). The messages also include process words (e.g., have, ok, and submit). Process words are associated with messages like "do you have," or "ok," or "submit your answer." As the task progresses, messages contain a smaller proportion of process words. The correlation between the trial variable and the proportion of process words equals -.540. As the task progresses, the messages include a larger proportion of names or labels like the descriptions and names in Figure 2. The correlation between the trial variable and the proportion of names or labels equals .552

6. For example, we calculate the difference between the regression coefficients for the FC and CB teams ($\theta_{FC} - \theta_{CB}$). The difference, along with the standard errors of the regression coefficients are used to calculate a z-statistic. We report differences between coefficients that are significant at the 0.05-level.

7. Our shared language variable considers the level of similarity between consecutive messages. The conversations on our teams overlap, which allows for a team shared language to emerge from a collective process. We could focus on the level of similarity between connected members of a team at the end of each trial (Lix et al., 2022). We calculate "discursive similarity" and add it to our SEM. We reach the same substantive conclusions for our shared language variable. The level of the discursive similarity variable declines as the task progresses, with the biggest decline occurring in the wheel network. Discursive similarity affects team accuracy. Except for the highest levels of discursive similarity, where the effects are equal, the largest effect is observed in the wheel network, while the smallest effect is observed in the FC network. It is not clear how we should interpret the results for the discursive similarity variable. The measure is calculated for observed network connections at the end of a trial, independent of how much or when two team members communicate on a trial.

8. CB teams realize the same coordination benefits as FC teams despite having fewer network connections. The shared language variable focuses on one kind of coordination. There is more to coordination than a shared language. During each trial, CB team members spend an average of 4.4 min sharing messages and share a total of 70.6 messages while communicating, for a message rate of 17.5 messages per minute. FC team members communicate for 5.3 min and share

94.1 messages, for a 21.4 message rate. Differences in message rates could affect team accuracy. A full treatment of the message rate variable is beyond the scope of the current manuscript. However, as a robustness check, the message rate variable was added to the structural equation model. A team's message rate improves with experience and at a faster rate for the FC teams. However, an increase in a team's message rate is more beneficial for team accuracy for the CB teams.

## References

Anderson BS, Butts C and Carley K (1999) The interaction of size and density with graph-level indices. *Social Networks* 21(3): 239–267.

Argote L and Ingram P (2000) Knowledge transfer: a basis for competitive advantage in firms. *Organizational Behavior and Human Decision Processes* 82(1): 150–169.

Argote L, Aven BL and Kush J (2018. The effects of communication networks and turnover on transactive memory and group performance. *Organization Science*, 29(2): 191–206.

Balkundi P and Harrison DA (2006) Ties, leaders, and time in teams: strong inference about network structure's effects on team viability and performance. *Academy of Management Journal* 49(1): 49–68.

Barber SJ, Harris CB and Rajaram S (2015) Why two heads apart are better than two heads together: multiple mechanisms underlie the collaborative inhibition effect in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41(2): 559.

Baronchelli A (2016) A gentle introduction to the minimal Naming Game. *Belgian Journal of Linguistics* 30(1): 171–192.

Bavelas A (1948) A mathematical model for group structures. Applied anthropology 7(3): 16–30.

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

Bojanowski P, Grave E, Joulin A, et al. (2017) Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics* 5: 135–146.

Brass DJ, Galaskiewicz J, Greve HR, et al. (2004) Taking stock of networks and organizations: a multilevel perspective. *Academy of Management Journal* 47(6): 795–817.

Burgess RL (1968) Communication networks: An experimental reevaluation. *Journal of Experimental Social Psychology* 4(3): 324–337.

Burt RS (1987) Social contagion and innovation: cohesion versus structural equivalence. *American Journal of Sociology* 92(6): 1287–1335.

Burt RS (2000) The network structure of social capital *Research in Organizational Behavior* 22: 345–423.

Burt RS (2005) *Brokerage and Closure*. New York: Oxford University Press.

Burt RS (2021) Structural holes capstone, cautions, and enthusiasms. *In: Small ML, Perry BL, Pescosolido B, and Smith N*

(eds). *Personal Networks: Classic Readings and New Perspectives*. New York: Cambridge University Press, 384–416.

Burt RS and Reagans RE (2022) Team talk: learning, jargon, and structure versus the pulse of the network. *Social Networks* 70: 375–392.

Butts CT (2006) Exact bounds for degree centralization. *Social Networks* 28(4): 283–296.

Centola D (2018) *How behavior spreads: The science of complex contagions*. Princeton, NJ: Princeton University Press, vol. 3.

Centola D and Baronchelli A (2015) The spontaneous emergence of conventions: an experimental study of cultural evolution. *Proceedings of the National Academy of Sciences* 112(7): 1989–1994.

Christie LS, Luce RD and Macy J (1952) Communication and learning in task-oriented groups. *Technical Report No. 231, Research Laboratory of Electronics, MIT*.

Clark HH and Wilkes-Gibbs D (1986) Referring as a collaborative process. *Cognition* 22(1): 1–39.

Csaszar FA and Siggelkow N (2010) How much to copy? Determinants of effective imitation breadth. *Organization Science* 21(3): 661–676.

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

Denrell J, Fang C and Levinthal DA (2004) From T-mazes to labyrinths: Learning from model-based feedback. *Management Science* 50(10): 1366–1378.

Fang C, Lee J and Schilling MA (2010) Balancing exploration and exploitation through structural design: the isolation of subgroups and organizational learning. *Organization Science* 21(3): 625–642.

Firth JR (1957) *A Synopsis of Linguistic Theory, 1930-1955*. Oxford: Studies In Linguistic Analysis.

Fraley C and Raftery AE (2002) Model-based clustering, discriminant analysis, and density estimation. *Journal of the American Statistical Association* 97(458): 611–631.

Freeman LC (1977) *A Set of Measures of Centrality Based on Betweenness*. Sociometry, 35–41.

Goldberg A, Srivastava SB, Manian VG, et al. (2016) Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *American Sociological Review* 81(6): 1190–1222.

Guilbeault D, Baronchelli A and Centola D (2021) Experimental evidence for scale- induced category convergence across populations. *Nature Communications* 12(1): 1–7.

Hong L and Page SE (2004) Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences* 101(46): 16385–16389.

Huang S and Cummings JN (2011) When critical knowledge is most critical: centralization in knowledge-intensive teams. *Small Group Research* 42(6): 669–699.

Huckman RS, Staats BR and Upton DM (2009) Team familiarity, role experience, and performance: evidence from Indian software services. *Management Science* 55(1): 85–100.

Katz N, Lazer D, Arrow H, et al. (2004) Network theory and small groups. *Small Group Research* 35(3): 307–332.

Knudsen T and Srikanth K (2014) Coordinated exploration: organizing joint search by multiple specialists to overcome mutual confusion and joint myopia. *Administrative Science Quarterly* 59(3) 409–441.

Koçak Ö and Warglien M (2020) When three's a crowd: how relational structure and social history shape organizational codes in triads. *Journal of Organization Design* 9: 1–27.

Krackhardt D (1999) The ties that torture: simmelian tie analysis in organizations. *Research in the Sociology of Organizations* 16(1): 183–210.

Krauss RM and Fussell SR (1991) Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition* 9(1): 2–24.

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

Le Q and Mikolov T (2014) Distributed representations of sentences and documents. *International Conference on Machine Learning*. PMLR, 1188–1196.

Leavitt HJ (1949) *Some Effects of Certain Communication Patterns upon Group Performance. Doctoral Dissertation*. Boston, MA: Massachusetts Institute of Technology.

Leavitt HJ (1951) Some effects of certain communication patterns on group performance. *Journal of Abnormal and Social Psychology* 46(1): 38–50.

Levy O and Goldberg Y (2014) Neural word embedding as implicit matrix factorization. *Advances in Neural Information Processing Systems*. MIT Press, 27, 2177–2185.

Lewis K (2003) Measuring transactive memory systems in the field: scale development and validation. *Journal of Applied Psychology* 88(4): 587.

Lix K, Goldberg A, Srivastava SB, et al. (2022) Aligning differences: discursive diversity and team performance. *Management Science* 68(11): 8430–8448.

March JG (1991) Exploration and exploitation in organizational learning. *Organization Science* 2(1): 71–87.

Maroulis S, Diermeier D and Nisar MA (2020) Discovery, dissemination, and information diversity in networked groups. *Social Networks* 61: 67–77.

Mathieu J, Maynard MT, Rapp T, et al. (2008) Team effectiveness 1997- 2007: a review of recent advancements and a glimpse into the future. *Journal of management* 34(3): 410–476.

Mehra A, Dixon AL, Brass DJ, et al. (2006) The social network ties of group leaders: implications for group performance and leader reputation. *Organization Science* 17(1): 64–79.

Mikolov T, Chen K, Corrado G, et al. (2013) Efficient estimation of word representations in vector space. *arXiv preprint arXiv: 1301.3781*.

Mora-Cantallops M and Sicilia MÁ (2019) Team efficiency and network structure: the case of professional league of legends. *Social Networks* 58: 105–115.

Moreland RL, Argote L and Krishnan R (1996) Socially Shared Cognition at Work: Transactive Memory and Group Performance. In: Nye JL and Brower AM (eds). *What's Social about Social Cognition? Research on Socially Shared Cognition in Small Groups*. Sage Publications.

Mukherjee S (2016) Leadership network and team performance in interactive contests. *Social Networks* 47: 85–92.

Mundlak Y (1978) On the pooling of time series and cross section data. *Econometrica: Journal of the Econometric Society* 46: 69–85.

Page SE (2007) Making the difference: applying a logic of diversity. *Academy Of Management Perspectives* 21(4): 6–20.

Park S, Grosser TJ, Roebuck AA, et al. (2020) Understanding work teams from a network perspective: a review and future research directions. *Journal of Management* 46(6): 1002–1028.

Rajaram S and Pereira-Pasarin LP (2010) Collaborative memory: cognitive research and theory. *Perspectives on Psychological Science* 5(6): 649–663.

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

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

Reagans R, Zuckerman E and McEvily B (2004) How to make the team: social networks vs. demography as criteria for designing effective teams. *Administrative Science Quarterly* 49(1): 101–133.

Reagans R, Argote L and Brooks D (2005) Individual experience and experience working together: predicting learning rates from knowing who knows what and knowing how to work together. *Management Science* 51(6): 869–881.

Reagans R, Singh PV and Krishnan R (2015) Forgotten third parties: analyzing the contingent association between unshared third parties, knowledge overlap, and knowledge transfer relationships with outsiders. *Organization Science* 26(5): 1400–1414.

Reagans R, Miron-Spektor E and Argote L (2016) Knowledge utilization, coordination, and team performance. *Organization Science* 27(5): 1108–1124.

Ren Y and Argote L (2011) Transactive memory systems 1985–2010: an integrative framework of key dimensions, antecedents, and consequences. *The Academy of Management Annals* 5(1): 189–229.

Rivera MT, Soderstrom SB and Uzzi B (2010) Dynamics of dyads in social networks: assortative, relational, and proximity mechanisms. *Annual Review of Sociology* 36: 91–115.

Rogge GO (1953) Personality factors and their influence on group behavior: a questionnaire study. *Technical Report No. 265, Research Laboratory of Electronics, MIT*.

Schilling MA and Fang C (2014) When hubs forget, lie, and play favorites: Interpersonal network structure, information distortion, and organizational learning. *Strategic Management Journal* 35(7): 974–994.

Selten R and Warglien M (2007) The emergence of simple languages in an experimental coordination game. *Proceedings of the National Academy of Sciences* 104(18): 7361–7366.

Shaw ME (1964) Communication networks. *Advances in Experimental Social Psychology* 1: 111–147, Academic Press

Sherf EN, Sinha R, Tangirala S, et al. (2018) Centralization of member voice in teams: its effects on expertise utilization and team performance. *Journal of Applied Psychology* 103(8): 813.

Simon HA (1973) The structure of ill structured problems. *Artificial Intelligence* 4(3-4): 181–201.

Smith SL (1950) *Communication Pattern and the Adaptability of Task-Oriented Groups: An Experimental Study*. Cambridge: Group Networks Laboratory, Research Laboratory of Electronics, Massachusetts Institute of Technology

Srivastava SB, Goldberg A, Manian VG, et al. (2018) Enculturation trajectories: language, cultural adaptation, and individual outcomes in organizations. *Management Science* 64(3): 1348–1364.

Sutton RS and Barto AG (1998) *Introduction to reinforcement learning*. MIT Press.

Ter Wal AL, Criscuolo P, McEvily B, et al. (2020) Dual networking: how collaborators network in their quest for innovation. *Administrative Science Quarterly* 65 (4): 887–930.

Tortoriello M and Krackhardt D (2010) Activating cross-boundary knowledge: the role of Simmelian ties in the generation of innovations. *Academy of Management Journal* 53(1): 167–181.

Tortoriello M, Reagans R and McEvily B (2012) Bridging the knowledge gap: the influence of strong ties, network cohesion, and network range on the transfer of knowledge between organizational units. *Organization Science* 23(4): 1024–1039.

Uzzi B (1999) Embeddedness in the making of financial capital: how social relations and networks benefit firms seeking financing. *American Sociological Review* 64: 481–505.

Weber RA and Camerer CF (2003) Cultural conflict and merger failure: an experimental approach. *Management Science* 49(4): 400–415.

Wooldridge JM (2010) *Econometric Analysis of Cross Section and Panel Data*. MIT press.