Centralization, language similarity, and performance: Renovating a classic experiment to identify network effects on team problem solving

January, 2021

Ray E. Reagans, Hagay C. Volvovsky, and Ronald S. Burt

Centralization, language similarity, and performance: Renovating a classic experiment to identify network effects on team problem solving

Abstract

Existing research illustrates a contingent association between a team's work assignment and the ideal network for superior performance. For a basic task, the ideal network is organized around a central individual. If the task is complex, the ideal network is decentralized and democratic. Language similarity is one reason why complex work requires a decentralized network. To perform well, team members must apply the same problem-solving framework, and decentralized teams have an advantage in reaching consensus. Recent research suggests that language similarity is more beneficial for performance when a network is centralized. The implications of this potential outcome are underappreciated. Even if centralized teams struggle to agree on what framework to use, if the performance implication of language similarity is larger in a centralized team, centralized teams could be preferable, even if the focal task is complex. We analyze the performance of seventy-seven teams working to identify abstract symbols, which is a complex task and which also requires language similarity. People are randomly assigned to different network conditions and work together for a number of trials. We find that language similarity improves with experience at a slower rate in centralized teams, but we also find that the language similarity effect on team performance is larger for centralized teams, large enough to shift the overall advantage to centralized teams. We also estimate the performance of teams working in networks that combine elements of centralized and decentralized networks. Performance is higher in teams that combine both network features.

INTRODUCTION

A key challenge facing a manager is evaluating a team's capacity for coordination and ultimately performance. A team's capacity for coordination depends, in part, on how team members are connected to each other (Balkundi and Harrison, 2006). The connections define a team's network. Existing theoretical arguments describe a basic contingency between how team members should be connected and features of their work assignment (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). If a team is solving a complex problem, team members are unlikely to know how to combine their individual efforts. Network connections provide team members with an opportunity to problem solve together until they find a valuable combination of their individual efforts (Argote, Aven, and Kush, 2018). For a complex problem, coordination and team performance improves when everyone on the team is connected. The ideal team network is decentralized. In a decentralized network, every team member is connected and can communicate directly. If a team is solving a basic problem, team members are more likely to know how their efforts fit together, and a single individual can coordinate their collective activities. The ideal network is organized around one individual or a team leader. The best team network is centralized.

The network contingency described above started to emerge in research efforts that began at MIT after World War II. The network imagery came from Bavelas (1948, 1950), the coordination experiment from Smith (1950) and the first published results from Leavitt's dissertation (1949, 1951). During a forty-year period, the MIT project produced over one hundred articles, books, or technical reports (Hummon, Dorein, and Freeman, 1990). Shaw (1964) provides an extensive review of the findings from 36 of the initial MIT experiments. The experiments include communication networks that range from being organized around a single individual to being fully connected. The teams perform a variety of tasks, including basic symbol-identification tasks

and more complex decision-making tasks (e.g., math problems, sentence completion problems, discussion problems). Consistent differences are observed between centralized and decentralized networks. When teams are solving a basic problem, team performance is higher in a centralized network, but when the teams are solving more complex problems, the pattern is reversed (Shaw 1964:122-124). Team performance is higher in a decentralized team network.

The initial research also suggests that when a team is solving a complex problem, it is important for team members to share a problem-solving framework. Using communication data from two of the initial experiments, Christie, Luce, and Macy (1952:chap 5) consider the association between a team's communication network, the emergence of a shared language and team performance. The task is a symbol-identification task. Team members are given a set of marbles, with only one marble in common. The team's goal is to identify the common marble. The experiment lasts for multiple trials and during the initial trials, the marbles are one color. Single color marbles are easier to describe and to identify correctly. During the later trials, the marbles are multicolored. Multicolored marbles are harder to describe and team members are more likely to make a mistake. During the later trials, the task is more complex and performance should be higher for teams working in a decentralized network. The empirical analysis focuses on the association between a team's communication network, language similarity (i.e., the extent team members use similar words and phrases to describe the multicolored marbles), and the effect language similarity has on performance. The empirical results do not include statistical tests and are tentative. However, the findings suggest sharing a language is more likely in a decentralized network and sharing a language increases the likelihood of identifying the common marble. Language similarity appears to mediate the association between a team's network structure and its performance. When team members are performing a complex task, they need to agree on their problem-solving framework, and decentralized teams have an advantage in reaching agreement.

The focus on language similarity in the MIT project anticipates more recent research emphasizing the importance of language similarity in shaping individual performance (Srivastava et al., 2018), including the potential for an individual's network to shape the language similarity effect on individual attainment (Goldberg et al., 2016). The network data comes from email exchanges and language similarity increases when the words in incoming and outgoing email messages overlap. The research findings indicate the language similarity effect on individual performance varies with the structure of the network surrounding the focal individual. If the individual occupies a central position in his or her network (i.e., the contact's in the individual's network are disconnected), language similarity has a positive effect on performance. When network members speak the same language, the central member is more likely to receive a favorable performance review and is less likely to involuntarily leave the organization.

The focus on language similarity can advance our understanding of an important team process, but also raises an important question. In particular, the focus on language similarity for complex work reintroduces the question of which team network is preferable. The previous discussion suggests an increase in centralization will have conflicting implications for performance. On the one hand, an increase in centralization reduces the likelihood of a team developing a shared language, and given the positive effect language similarity can have on team performance, the increase should reduce team performance. On the other hand, an increase in centralization can strengthen the positive association between language similarity and team performance, which should enhance performance. Even if centralization reduces the likelihood of a team developing a shared language, if centralization also increases the performance implications of a shared language, the overall effect on team performance is unknown. Depending on the relative magnitude of the two performance pathways, a decentralized or centralized team could have

the advantage. To answer this question, we must know how much centralization affects language similarity and also how much centralization shapes the performance implications of a shared language.

We attempt to shed light on this important question by analyzing the association between a team's communication network and language similarity also how much a team's network moderates the association between language similarity and team performance. We use a symbol identification task (Clark and Wilkes-Gibbs, 1986:11; Selten and Warglien, 2007). Instead of marbles, we ask our teams to identify abstract symbols. An individual could describe the abstract symbols with a large number of words and phrases. To do well, team members must use the same terms and phrases. The task allows us to analyze how a team's network affects how quickly team members reach agreement and to also estimate how much agreement affects team performance.

CENTRALIZATION AND TEAM PERFORMANCE

Figure 1 illustrates two pathways from centralization to team performance. Consistent with insights from the MIT project, language similarity mediates the association between centralization and team performance. Consistent with current research, centralization also moderates the association between language similarity and performance. The path from language similarity to performance highlights the importance of coordination for performance. When individuals are assigned a complex problem, it is unclear how they should combine their individual efforts. Effective coordination is unprogrammed (March and Simon, 1958) or relational (Hoffer Gittell, 2002). Unprogrammed or relational coordination often forms and adapts as the team performs it tasks (Faraj and Sproull, 2000). As team members work together, they develop their own language and shorthand terms to describe element of their work context (Weber and Camerer, 2003; Burt and Reagans, 2020). A shared language

facilitates communication (Krauss and Fussell, 1990), which enables team members to coordinate their activities more effectively (Clark and Marshall, 1981) and improves their performance (Weber and Camerer, 2003; Reagans, Miron-Spektor, and Argote, 2016).

——— Figure 1 About Here ———

We lack systematic evidence linking a team's communication network to language similarity. There is, however, research describing the effect language similarity can have on the formation of individual relationships (Kovacs and Kleinbaum, 2020). When two individuals speak the same language, they are more likely to develop a relationship and once a network connection is established, interpersonal dynamics increase the level of language similarity in the on-going relationship. While we do not have systematic evidence linking network structure to language similarity, we do have results from related research. Language similarity is a form of consensus. There is a large body of research documenting the effect interpersonal relationships can have on consensus (Centola, 2020:11-62 provides an extensive review). Consensus is more likely when two individuals communicate with the same people and those people also communicate with each other (Burt, 1987). Research on political polarization is a vivid example of how network connections can shape individual behaviors and beliefs (DellaPosta, Shi, and Macy, 2015).

Given prior research on networks and consensus, we expect for language similarity to emerge more quickly in a decentralized network. In a decentralized network, there are more direct and indirect communication channels, increasing the potential for interpersonal influence. And more interpersonal influence should translate into a higher level of language similarity (Saint-Charles and Mongeau, 2018). Evidence on problem solving in a group leads to a similar prediction. When a group of individuals problem solve together, they influence each other's choices and decisions, narrowing the range of choices and alternatives they consider (Rajaram and Pereira-

Pasarin, 2010; Barber et al., 2015). Teams that focus their problem-solving efforts on a smaller number of alternatives are more likely to reach consensus.

Our teams work together for a number of trials. Language similarity can be expected to increase with experience (Clark and Wilkes-Gibbs, 1986). Our current discussion has focused on how the team's network affects the rate at which experience translates into consensus. We expect for language similarity to increase with experience at a slower rate for a centralized team. If a team's network is centralized, communication flows through a single individual. Team members have fewer opportunities to influence each other's word choices, and as a result should take relatively more time to reach consensus.

——— Figure 2 About Here ———

We use four networks in our experiment. The networks are displayed in Figure 2. The CLIQUE network is in the lower left of the figure. The CLIQUE is a closed network in which everyone is connected to everyone else. The CLIQUE is decentralized. The WHEEL network is in the lower right. The WHEEL network contains a single leader. The WHEEL is centralized. We discuss the top networks in the next section. Since our centralized network is a WHEEL and our decentralized network is a CLIQUE, we test the following hypothesis.

<u>Prediction 1a: The positive effect experience has on language similarity will be less positive in a WHEEL network than in a CLIQUE network.</u>

The previous discussion describes the tendency for language similarity to emerge more slowly on centralized team. There are reasons to suspect the value of language similarity will also vary across the network conditions. Once again, the research on problem solving in groups informs our thinking. Individuals who problem solve alone do so free from the choices and decisions

made by each other, and as a result will collectively entertain a wider array of choices and decisions. The autonomy can be beneficial. Individual problem solvers produce a larger number of creative ideas than the same number of individuals working together (Taylor, Berry, and Block, 1958; Diehl and Stroebe, 1987; Putman and Paulus, 2009). The autonomy can be especially beneficial if the problem the individuals are solving shifts (Shore, Bernstein, and Yang, 2020).

We expect for centralized teams to take longer to reach consensus because team members have fewer opportunities to influence each other's word choices. But at the same time, members of a centralized team should entertain and consider a wider array of word choices. Considering a wider array of word choices could be beneficial. Teams that consider a wider array of words and phrases, could ultimately agree to use words which allow for more effective coordination. This line of thinking suggests that the performance implications of language similarity will be higher in a centralized team. While we do not know the answer to this question for teams, as we noted earlier, there is evidence for this effect for individuals. Language similarity has a positive effect on individual performance, when an individual occupies a central position in his or her communication network. Given this finding and the discussion above, we test the following prediction.

Prediction 2a: The positive language similarity effect on team performance is more positive in a WHEEL than in a CLIQUE network.

It is important to distinguish our prediction from previous research. In particular, Goldberg, Srivastava, and Manian (2016:1207) found that language similarity reduces individual performance, when an individual works in a constrained communication network. A constrained network is decentralized. We do not expect for language similarity to have a positive effect for

centralized teams and a negative effect for decentralized teams. We expect for language similarity to have a positive effect for centralized and decentralized teams. Our symbol identification task requires coordination. And language similarity improves coordination (Reagans et al., 2016; Lix et al., 2020). We simply expect for the positive language similarity effect on performance to be more positive for centralized teams. Our prediction is based on the idea that there are more effective ways to describe the abstract symbols. Teams that are characterized by less interpersonal influence have more opportunities to find effective words and phrases, and if they decide to use what they find, team performance will improve.

Networks with Two Brokers

We have focused our discussion on a team either having a centralized or a decentralized network, but one can imagine combining features of the two networks to realize the benefits both networks provide. On the one hand, redundant communication channels in a CLIQUE network should allow language similarity to increase faster but on the other hand, non-redundant relationships in the WHEEL has the potential to identify more effective words and phrases. The WHEEL network could be modified to include two central individuals instead of one (Rogge, 1953:18). Both individuals would be responsible for helping the team to identify and develop common words and phrases, increasing the potential for consensus on the team. With limits on team size, introducing an additional central individual would come at the expense of a distinct perspective on the team. The potential for a centralized team to identify higher quality word choices would be diminished but not completely offset.

The central individual in the WHEEL is commonly referred to as a broker. We consider two versions of the network with two brokers. 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 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 CB network is in the upper left panel of Figure 2 and the DB network is in the upper right panel.

In the CB team, team members should be able to consider a wider array of words and phrases before ultimately deciding on what words and phrases to use. For the benefits the CB network represents to be realized, the two central individuals must learn how to work together (Ter Wal et al., 2020). We do not know how much experience the two central individuals will need before they learn how to work together effectively.

The DB team also reduces the team's dependence on a single individual. The additional broker creates indirect communication channels between the team members. The channels are indirect but introduce redundant communication which could increase the team's capacity for reaching consensus. As with the CB network, the additional broker reduces but does not eliminate the number of independent perspectives on the team. The relationship between two brokers is the difference between the CB and DB teams. The DB teams provide a more natural baseline for evaluating the potential value created by the relationship between the two brokers. We do not know when the two brokers will learn how to coordinate their efforts, but we expect for individuals in a CB network to reach consensus faster than the individuals working in a DB network. To describe the relative merits of the CB and the DB networks, we can apply the same rationale we used to distinguish the CLIQUE and the WHEEL networks. CB teams should have the advantage in reaching consensus, but the DB teams should have the edge in terms of the performance implications of consensus.

Prediction 1b: The positive effect experience has on language similarity is more positive in a CB network than in a DB network.

Prediction 2b: The positive language similarity effect on team performance is more positive in a DB network than in a CB network.

We can also compare the two broker networks to the CLIQUE and the WHEEL networks. The CB network lies between the CLIQUE and the WHEEL networks. We can create the CB network by either adding three connections to the WHEEL network or by removing three connections from the CLIQUE. The similarity outcomes we observe in the CB network, both the experience effect on language similarity and the magnitude of the language similarity effect on team performance, should lie between the outcomes we observe in the WHEEL and CLIQUE networks. For example, we expect for language similarity to increase with experience at a higher rate in the CB network than the WHEEL network, but we also expect for the magnitude of the language similarity effect on team performance will be higher in the WHEEL than in the CB network. The same logic applies to the DB network. We can create the DB network by changing four relationships in either the CLIQUE or the WHEEL network. The WHEEL and the CLIQUE should bound the outcomes we observe in the DB network. For example, there is more redundancy in the DB network than the WHEEL, so we expect for language similarity to increase faster with experience in the DB network than in the WHEEL.

Experimental Design

We modify the original MIT experimental design to operate in an online setting. In the original experiment, individuals work at a table passing written notes under screen partitions (Leavitt, 1951:41). We replace the table with a computer interface. Putting aside people involved in pre-testing software and protocol, the final subjects are 385 men and women in 77 teams using the subject pools of MIT's Behavioral Research Lab (25 teams) and Harvard Business School's

Computer Lab for Experimental Research (52 teams). The subject pool contains students from MIT, Harvard, and neighboring schools, along with individuals from the surrounding communities. Pre-testing indicates that non-native English speakers find the learning task difficult to complete with native speakers, and older subjects often have difficulty with the chat-boxes and other features of the software, so participation is limited to native English speakers and people between the ages of 18 and 55.

Qualified subjects are assembled at a laboratory, take a seat at one of the separated computer cubicles, complete initial questions, and are told the experiment involves playing 15 trials of a team coordination game with structured communication between players. Each subject is in an individual cubicle. All communication is through the subject's computer. Each team is assigned a network at random. Subjects are randomly assigned to teams and to a network position on each team. To maintain balance across network conditions, if a network structure has been infrequently assigned, a team or two toward the end of the day is assigned manually to the network. Subject assignment to network position is at all times random. Subjects occupy their assigned network position in all trials of play.

Figure 3 shows the subject's screen. The upper-left of the screen shows a set of five symbols assigned to the subject. This is the subject's card or "hand." The five symbols in each subject's hand come from a superset of the six symbols in Figure 4. There are six combinations of the Figure 4 symbols taken five at a time. One symbol is shared in five of the possible six combinations. That is sufficient for each subject on a team to receive a different hand, with one symbol shared by all team members. The identification of the symbols in brackets is presented in Figure 4 to facilitate discussion, but you can see in Figure 3 that the symbols are presented to subjects without any short-cut identification. More often than not, teams correctly identify the symbol they share (brackets in Figure 4) — with the statistically significant exception of symbol

E, which is identified correctly in exactly half of the games in which it is the shared symbol. The symbols are so called “tangrams,” which originated in China many hundreds of years ago, became popular in Europe in the 19th century, then spread again during World War I. Several thousand tangrams can be constructed from the seven generative shapes (see Wikipedia “tangram” for general background). Tangrams have been useful in teaching geometric concepts and studying language, the latter because people have to create language distinguishing the odd symbols in order to coordinate with one another about the symbols. The six in Figure 4 are taken from a well-known study of subjects creating language to coordinate tangram sequences (Clark and Wilkes-Gibbs, 1986:11).

——— Figure 3 and Figure 4 About Here ———

To learn what other people have in their hands, subjects communicate by clicking on a teammate in the dialogue box at the top of the screen (“A” in Figure 3). Teammates listed in the dialogue box are the ones with whom a subject is allowed to communicate. The screen in Figure 3 is for player 2, who has access to all four teammates — indicated by options in the dialogue box for communication with player 1, 3, 4, or 5. As a subject communicates with teammates, a teammate-specific dialogue box at the bottom of the screen accumulates exchanges (“B” in Figure 3). The messages sent and received during a trial can be reviewed by moving the dialogue-box slider up or down.

Subjects talk to one another about the tangrams in their hands until making a guess about what tangram they have in common. 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 (“C” in Figure 3). Dots at the top of the screen darken as teammates submit answers, so there is some social pressure on a subject to submit an answer as others have already done so (“D” in Figure 3). Subjects do not see teammate answers, but they see from the darkened dots how many have submitted answers. An option was provided for subjects to “reconsider” their

submitted answer. When a subject submits an answer, the “submit answer” button on the screen (“C” in Figure 3) turns into “reconsider.” If “reconsider” is clicked before all other teammates submit answers, the number of darkened dots decreases by one and the trial stays open until the subject and all teammates submit an answer. 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 hand, and the next trial begins. Teams were given 75 minutes to finish all fifteen trials.

Team Messages

Our argument for language similarity implies language specialization. We assume that with experience problem-solving together, team members will eventually agree to use a smaller set of words and phrases, which should allow for more efficient and effective communication. Two indicators of efficient communication are the number of messages team members share and the number of words in each message. Figure 5 shows the rate at which messages and words per message decrease across the 15 trials. To complete their first trial together, teammates exchange an average of 147 messages (Figure 5A). The messages average seven words in length (Figure 5B). In the 15th trial, the average team uses fewer and shorter messages (respective averages of 37 messages and four words per message). The transition to more efficient messaging, obvious in Figure 5, is statistically significant and robust to a variety of controls (Burt et al., 2020:Table 8 ff.).

In our setting, language specialization is more than a shift to a smaller subset of words and phrases. With experience, team members move from attempting to describe the symbols using phrases and sentences to either describing the symbols in terms of their perceived actions or physical properties; or to simply naming or labeling the shapes; labels such as “priest”, “angel”, “absangel,” “sittingman,” “sitkick,” or “wallflowerkicker.” The labels are team jargon words (Burt and Reagans, 2021). The shift can be illustrated by considering the composition of team messages with respect to parts of speech. There is a general distinction in language between function and content words. Function words indicate relations between content words in a sentence. Example function words are pronouns (he is a new victim), prepositions (go to the store), articles (a, the), and auxiliary verbs (verbs that indicate the tense, mood, or voice of other verbs, e.g., I would have gone). Function words are often described as the glue that holds a sentence together. Content words are sentence elements with clear meaning that are held together by function words. For example, nouns (he is a new victim), verbs (he is a new victim), and adjectives (he is a new victim).

——— Figure 5 About Here ———

With the exception of personal pronouns, the parts of speech for function words are fixed. The parts of speech for content words vary. Context is important for defining parts of speech. A content word can be an adjective in one message and an adverb in another message, all depending on how the word is used. We assign parts of speech to the words in our messages using spaCy. spaCy is an open source library for Natural Language Processing (NLP) in Python. spaCy assigns parts of speech using Universal tags (e.g., AUX = auxiliary, ADV= adverb, PRON = pronoun). For each message, we create a count for each tag and we divide those counts by the number of words in a message to define proportions for the different parts of speech. We highlight three parts of speech: function words; nouns, proper nouns, and adjectives; and verbs and adverbs. Figure 5D shows the trial-by-trial decrease in function

words. Function words are 43% of the words on trial one of the messages; 36% of the words on trial five; but only 21% of the words on the final trial. Simultaneously, Figure 5C shows a steady use of verbs and adverbs. Verbs and adverbs represent 19% of the words on trial one; 20% on trial five; and 24% on the final trial. There is a sharp increase in the use nouns and adjectives. Nouns and adjectives represent 20% of the words on trial one; 28% on trial five; and 41% on the final trial. The symbol names (e.g., “priest”, “angel”, “absangel”) described above are nouns.

The decline in function words is surprising. Our symbol-identification task requires coordination. Prior research has emphasized the importance of function words in producing language-based coordination. Examples include success in negotiations (Taylor and Thomas, 2008; Bayram and Ta, 2018) and social attachment and cohesion in student teams (Gonzales et al., 2010; Tausczik and Pennebaker, 2013). Jargon words are content words. As team members shift to using jargon words to coordinate their choices, the use of function words declines. The two outcomes are interdependent.

We use recent developments in NLP to illustrate the substance and general meaning of the words our subjects use. 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 our message data. fasttext is an open-source library for text representation and classification. We follow convention in exclude

stopwords from our text analysis. Stopwords are words which occur more frequently in text and so have ambiguous meaning. Some stopwords are function words (e.g., “the”, “a”, “in”) while some are content words (e.g. “down”, “above”, “after”).

To illustrate the broad meaning of words, we cluster analyzed our word vectors to identify words with similar meanings (Fraley and Raftery, 2002).[1] The cluster analysis indicates there are five broad word categories in the message data. Table 1 contains information on the content of the five clusters. For each cluster, we bold three representative words, the total number of words in each cluster and how frequently individuals use words in each cluster. We refer to the words in column I as process words. Process words include words for selecting and confirming choices (“okay so not that one lol”, “yes I have sitting man”, “okay three crossed off, three to go”). The process cluster also includes words from commands and questions (“let’s go for it”, “ok everyone submit”, “who submitted the answer?”, “ask about buddha hurry”). Columns 2 through 4 contain words individuals generally use to describe the symbols. Individuals use the words in Column 2 the most frequently. The words indicate attempts to name the symbols (“I have abs man, rabbit, arms in the air, falling backpack”). The words in Column 3 are more action oriented and generally refer to what the symbols are doing (“my sitting man is facing left”). Cluster 3 also includes actions that were part of team process. Column 4 contains words which refer to the shape or general features of the symbols (“one with a nose facing left (a half trapezoid)”). The words in column 5 are more miscellaneous. Individuals use the words infrequently and often use the words a variety of ways. For example, in one message “ya” appears to mean yeah and in

[1] We cluster analyze our word vectors using a gaussian mixture model (gmm) (Fraley and Raftery, 2002). A gmm attempts to describe input data as a mixture of multi-dimensional gaussian probability distributions. Each mixture is a distribution and each data point is assumed to be drawn from one mixture. The method can be used to find clusters in data, with each cluster representing a mixture. The model has a number of virtues, including allowing us to assess overall model fit and how much a word fits with its assigned cluster.

another “ya” means yay. The teams appear to use different approaches to identify the symbols. The words in Cluster 2 are names or labels. The words in Cluster 3 focus on action and movement and the words in Cluster 4 describe the symbols in terms of their features and shapes.

——— Table 1 About Here ———

Figure 5D illustrates how the relative frequencies of the word categories, the substance of the messages, vary across the trials. A number of patterns are noteworthy. First, the individuals use the miscellaneous (the squares), feature (the circles) and action (the diamonds) words less frequently than they use process and name words. Second, as experience increases, process words occur less frequently, while words from the name category (i.e., the “labels” or jargon) occur more frequently. The decline in process words indicate team members are becoming more efficient in their interactions. The increase in words from the name category indicates the individuals believe focusing on labels is a more effective way to identify the symbols. Team members believe using labels is more effective than focusing on what the shapes appear to be doing or how they look.

Language similarity

To measure the level of language similarity between messages, we first represent each message as a vector by taking the average of the word vectors it contains (Le and Mikolov, 2014).[2] We then use cosine similarity to measure the level of language similarity between consecutive messages on a team:

$$\text{language similarity} = \frac{M_t M'_{t+1}}{\sqrt{M_t M'_t}\sqrt{(M_{t+1} M'_{t+1})}}$$

[2] 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 attempts to include this additional information in our message vectors (Le and Mikolov, 2014).

For each team-trial, we average the language similarity scores to measure the extent team members communicate with similar words and phrases. Figure 5E shows how the average language similarity scores vary with team experience. As the number of trials increases, language similarity increases but the increase is not uniform. Language similarity improves slowly during the initial trials. Language similarity increases steadily from the fifth to approximately the tenth trial. But during the final trials, there is more variation from trial to trial. There is a decline in language similarity on the thirteenth and fifteenth trials of the experiment. Time is often running short during the final trials of the experiment and some teams could find it more difficult to stay on task and communicate effectively.

We also calculate language similarity using only content words. The language similarity variable based on content words alone has virtually the same association with the trials but the variable is lower initially. The level difference suggests some of the function words are important. Numbers are function words and appear to be important. This isn't to suggest the team members are coordinating on numbers. Team members often use numbers to refer to the sequence of the symbols on their cards. They mistakenly believe there an association between the sequence of the symbols on their cards and the sequence of the symbols on the cards of other team members. When everyone is using numbers (the number 1 and the word one) in their messages, the level of semantic similarity is higher. One and 1 have similar meanings. For example, the correlation between the word vector for 1 and the word vector for one is .64. The correlation between the vector for number 2 and the vector for the word two is .85. While both correlations are positive, the more positive correlation for the number 2 and the word two indicates our subjects use them interchangeably.

Another way to think about language similarity is the consistent use of words. When a word is used in one context and not in other contexts, there is more agreement about its meaning.

Consistency in use illustrates clarity. A useful indicator of word clarity is a word's vector length or magnitude (Schakel and Wilson, 2015). For each message, we calculate the average clarity of function and content words. Clarity is higher for content than for function words (2.0 vs. 1.7). Content word clarity increases with experience (r = .43, p < .05), while there is no association between experience and function word clarity (r = -.05, p = .13). As teams gain experience, they use high clarity content words.

Overall, our analysis of the words and phrases our subjects use to coordinate their efforts indicates that language similarity emerges from a trial and error learning process where team members identify and even create content words they can use to coordinate their collective efforts. Words which allow team members to label the symbols are critical. The labels are more likely to be nouns, proper nouns, and adjectives. The words facilitate coordination because they have more clarity.

Estimation

To estimate our network effects, we need to address two estimation issues. First, teams complete the task multiple times. 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). The fixed effect controls for unobserved differences between teams. Second, we expect for the network conditions to affect language similarity and for language similarity to affect team performance. Language similarity 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 let the random effects be correlated with our predictors. The SEM framework also allows for us to estimate the path from our network conditions to language similarity and

from language similarity to team performance. Our indicator of team performance is the number of correct responses on a trial. Our performance variable is a count of the number of teammates who correctly guess the shared symbol on the focal trial. The dependent variable in our performance equation is assumed to follow a binomial distribution. Empirical results to be presented lead to the same substantive conclusions if our dependent variable is the proportion of correct responses.

We include in our equations a variable for team experience, which is the log of our trial variable. A large body of research has documented the positive effect team experience can have on team performance (Reagans et al., 2005; Huckman et al., 2009). We also include a number of control variables. Preliminary analysis indicates some symbols are harder to identify (Symbol E in Figure 4), so we control for the shared symbol on a trial. 77 teams start the experiment but 32 teams did not complete all fifteen trials. Teams vary in terms of when they collapse. No teams collapse during the first four trials. Four percent collapse between five and nine. The proportion jumps to seventeen percent on trial ten and declines to three percent during the final trials. We include two control variables to capture this dynamic. We include a dummy variable 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 exits early. Finally, our data comes from two labs, one at MIT and the other at Harvard. We do not expect for lab location to affect our results, but we include a binary variable set equal to one if the data was collected at MIT and zero if the data was collected at Harvard. Descriptive statistics for the variables we use in our analysis are in Table 2.

——— Table 2 About Here ———

Network Effects on Language Similarity

To test our first predictions, we regress language similarity on the log of our trial variable and include interaction terms between the trial variable and the network conditions. The WHEEL network is the excluded category. The results are in Column I and Column II of Table 3. The main effects are in Column I. Interaction terms are included in Column II. The regression coefficient for the trial variable is positive and significant ($b = .012$, $z = 3.52$, $p < 001$). The coefficient defines the rate language similarity increases with increasing experience. The coefficient defines an experience curve and the main effect defines the experience curve for the WHEEL. The significant interaction terms indicate the experience effect varies across the network conditions. The coefficients for the interaction terms are all positive and significant. Language similarity increases with experience at the lowest rate in the WHEEL network. The coefficients for the interaction terms are relative to the WHEEL. The estimates indicate the experience curve is steeper in the CB network (.029 slope-adjustment, $z = 4.33$, $p < .001$), followed by the CLIQUE network (.014 slope-adjustment, $z = 3.26$, $p < .001$) and then the DB network (.011 slope-adjustment, $z = 2.54$, $p < .05$).

——— Table 3 About Here ———

We use the MARGINS command in STATA to compare the magnitude of the experience effect in the different network conditions. The positive coefficient for the interaction between experience and the CLIQUE network condition indicates the experience curve in the CLIQUE is steeper than the curve in the WHEEL. Results from the MARGINS command indicate the experience curve in the CB network is steeper than the curve in the DB network ($\Delta = .009$, $z = 1.99$, $p < .05$). When we compare the CB network to the WHEEL and CLIQUE networks, the experience effects in the CB and CLIQUE networks are equal ($\Delta = .006$, $z = 1.20$, $p = .23$), while the experience effect in the CB network is larger than the effect in the WHEEL ($\Delta = .020$, $z =$

4.33, $p < .001$). When we compare the DB network to the WHEEL and CLIQUE networks, the experience effect is larger in the DB network than the experience effect in the WHEEL (Δ = .011, $z = 2.54$, $p < .05$), while the effects in the DB and CLIQUE networks are equal (Δ = -.004, $z = -.080$, $p = .43$).

In addition to comparing the slopes, it is informative to consider the levels of language similarity across the network conditions trial-by-trial. For example, the experience effects in the DB and CLIQUE networks are equal, which is surprising. But if we compare the observed levels of language similarity in the DB and CLIQUE networks trial-by-trial, a difference emerges after the sixth trial. The difference equals approximately -.015. The difference is only marginally significant but consistent with the prediction that language similarity would be higher in the CLIQUE than in the DB network.

If we step away from the details of the results, we find broad support for our first set of predictions. As we predicted, language similarity emerges faster in the CLIQUE network than in the WHEEL network. Language similarity also emerges faster in the CB network than in the DB network. When we compare the CB and DB networks to the CLIQUE and WHEEL networks, the results are less definitive but consistent with our predictions. We expected for the experience effects in the CB and DB networks to lie between the outcomes we observe in the CLIQUE and WHEEL networks. When we focus on the DB network, the experience curve is steeper than the curve we observe in the WHEEL and when we compare the DB network to the CLIQUE network trial-by-trial, we observe a marginally significant difference between the CLIQUE network and the DB network, with language similarity levels being higher in the CLIQUE network. When we focus on the CB network, the experience curve in the CB network is steeper than the

experience effect we observe in the WHEEL network. But the experience effects in the CB network and the CLIQUE network are equal.

We did not expect for the language similarity outcomes to be equal in the CB and CLIQUE network conditions. We can only speculate why. It is possible the average level in CLIQUE teams is lower than expected. With the benefit of hindsight, it is easier to appreciate that CLIQUE teams have features that can undermine communication. There are more relationships in the CLIQUE network and each interaction requires time and attention. When there are limits on network time and energy, there is less time to spend in each individual relationship. Moreover, since there are more relationships in the CLIQUE network, there is also a greater potential for variation in communication across those relationships. Unless everyone in a CLIQUE network is communicating with everyone else in the network at the same point in time (e.g., over email or in the same geographic location), which would reduce the CLIQUE network to a single communication channel; during the initial stages of a learning process, there is the potential for some miscommunication, which would reduce the observed level of language similarity.

The average level of language similarity in the CB network could be higher than we expected. We assumed it would take time for the two brokers to learn how to coordinate their behavior. They could have learned at a much faster rate. The higher language similarity levels in CB teams could also be a function of the disconnects between team members. While there are fewer channels in the CB network, the ones that do exist have more bandwidth. Since there are fewer relationships, the bandwidth in the remaining relationships can be higher and there are fewer opportunities for miscommunication (Aral and Van Alstyne, 2011). The holes between individuals in the CB networks act as buffers limiting communication and the higher bandwidth connections can provide a foundation for effective communication.

Language similarity effects on team accuracy

To test our second set of predictions, we regress team accuracy on language similarity and interact the language similarity variable and the network conditions. The interactions allow the language similarity effect on team accuracy to vary across the network conditions. The WHEEL network is the excluded category. The results are in Columns III and IV of Table 3. The dependent variable in Equation 1 is language similarity. The dependent variable in Equation 2 is team accuracy. We focus on the Equation 2 estimates in Column IV. The coefficient for language similarity is positive and significant ($b = 13.16$, $z = 6.52$, $p < .001$). There is one significant interaction. The interaction between language similarity and the CLIQUE network condition is negative and significant (-10.49, $z = -4.03$, $p < .001$). The estimates indicate the magnitude of the language similarity effect is larger in the WHEEL than the CLIQUE network. We use the MARGINS command in STATA to compare the magnitude of the language similarity effect between all of the network conditions. The WHEEL, DB, and CB networks are all broker networks. The magnitude of the language similarity effect on team accuracy does not vary across the WHEEL, DB, and CB networks. There is no difference between the CB and DB networks ($\Delta = 0.69$, $z = 0.37$, $p = .71$); the CB and the WHEEL networks ($\Delta = 1.12$, $z = 0.61$, $p = .54$); or the DB and the WHEEL networks ($\Delta = -1.81$, $z = -0.94$, $p = .35$). While we do not observe any differences between the WHEEL, DB, and CB networks, we do observe significant differences between the three and the CLIQUE network. When compared to the CLIQUE network, the language similarity effect on team accuracy is larger in the WHEEL ($\Delta = 6.97$, $z = 3.58$, $p < .001$), the CB ($\Delta = 5.84$, $z = 3.14$, $p < .05$) and the DB network ($\Delta = 5.14$, $z = 2.64$, $p < .05$).

The WHEEL, DB, and CB networks are all broker networks. To highlight the contrast between the CLIQUE network and the broker networks, we re-estimate the SEM equation and replace the network conditions with a binary variable equal to one for CLIQUE teams and zero for the broker networks. When we do, the coefficient for the language similarity variable is 12.09 ($z = 10.29$, $p < .001$) and the coefficient for the CLIQUE network interaction term is -9.40 ($z = -4.75$, $p < .001$).

We find partial support for our second set of predictions. As predicted, the language similarity effect in the WHEEL is larger than the effect in the CLIQUE network. However, the language similarity effect in the DB network is not larger than the effect in the CB network. The language similarity effect in the WHEEL network is not larger than the effects in the DB and CB networks, but the effects in the DB and CB networks are larger than the language similarity effect in the CLIQUE network. Technically, we observe language similarity effects on performance in the DB and CB networks that lie between the effects we observe in the WHEEL and CLIQUE networks. But the language similarity effect in the WHEEL network is not greater than the effects we observe in the DB and CB networks. The effects in the WHEEL, DB, and CB networks are equal.

The Best Team

The empirical results illustrate how our network conditions affect language similarity and also how the network conditions moderate the language similarity effect on team accuracy. The results indicate the team networks have distinct advantages. For example, teams working in the CLIQUE and CB networks have an edge in reaching consensus, while the teams working in the WHEEL, CB and DB networks have an advantage in translating consensus into accuracy. Language similarity varies across the network conditions but so does the language similarity effect on team performance. When we compare the teams in terms of overall accuracy, the CB

teams are the most accurate. They are more accurate than teams working in the WHEEL (Δ = .23, z = 9.78, p < .001), the DB (Δ = .26, z = 12.71, p < .001), and the CLIQUE (Δ = .58, z = 8.73, p < .001) networks. The WHEEL teams are next. Teams assigned to the WHEEL are more accurate than the teams assigned to the DB (Δ = .03, z = 2.21, p < .05) and the CLIQUE (Δ = .35, z = 5.13, p < .001) networks. Finally, the teams assigned to work in the DB network are more accurate than teams assigned to work in the CLIQUE network (Δ = .32, z = 4.72, p < .001). Thus, while individuals working in the WHEEL and DB teams struggle to reach consensus, the consensus they are able to reach has a much larger effect on team accuracy than the consensus reached in the CLIQUE teams. While the CLIQUE teams reach consensus faster, the consensus they reach is less beneficial for team performance. Like the CLIQUE teams, the CB teams are able to reach consensus, but like the WHEEL and DB teams, they benefit more from the consensus they reach.

Summary and Discussion

It is useful to quickly summarize our results. Our results are consistent with the framework initially described by Christie and his colleagues. Decentralized teams have an advantage in agreeing on what words and phrases to use and agreement improves team performance. We also find evidence for more recent research which notes the enhanced value of agreement if an individual occupies a central position a network. Agreement is more performance enhancing for our centralized teams. The two effects come together to inform our understanding of which teams have an advantage when they are solving a complex task. Teams working in a decentralized network have an advantage in reaching consensus, but the performance implications of reaching consensus are larger for teams working in a centralized network, shifting the overall advantage to teams working in a centralized network.

Our findings appear to call into question the idea that centralized teams are better for basic tasks, while decentralized teams are better for complex tasks. Our task is complex but our centralized teams do better. One could reasonably ask, just how generalizable are the results? We ask our teams to work together to identify abstract symbols, which is a coordination task. Our results indicate that even for this coordination task, creative problem solving is important. Our teams have to decide what words and phrases to use, including the words and phrases they create, to coordinate their efforts. In our setting, there are better and worse ways to coordinate. Deciding where to coordinate requires the consideration of a wider array of choices and decisions, and creative problem solving is a predictable end result of this process. The strong interpersonal influence which emerges in a decentralized network can be beneficial but can also put a limit on team performance. The redundant communication channels that exist in a decentralized network can improve coordination but can also result in team members coordinating their efforts on less productive choices and ideas.

Perhaps we can establish external validity for our results by comparing them to findings in the field. For example, the CB teams are the best teams in our study. The CB teams combine element of the CLIQUE and WHEEL networks. The results for our CB teams are consistent with research on the social capital of teams. In particular, in Reagans and Zuckerman (2001; Reagans, Zuckerman and McEvily, 2004), the ideal team network is a clique network internally, but contacts outside of the team are disconnected from each other. The team occupies a central position between contacts who are disconnected from each other. Disconnects in the team's external network increases the team's capacity for creative problem solving, while numerous relationships inside the team increase the team's capacity for collaboration and coordination. All of our team networks are internal and all of our NON-CLIQUE teams do better than CLIQUE teams. Our ideal network appears to be in conflict with the ideal network described by Reagans and Zuckerman and other scholars who study team network effects in the field.

The disagreement is surface level. In our CB network and in Reagans and Zuckerman's ideal network, team performance is enhanced when a team's network allows its members to identify and evaluate a wider array of choices and decisions, and to coordinate their efforts on more effective choices and decisions. We can also resolve the apparent conflict by noting a division of labor inside our teams. All team members are responsible for problem solving and coordinating their efforts, but central team members are especially critical for coordination. Central team members are more likely to be recognized as a team leader (Burt, Reagans, and Volvovsky, 2020). Leadership can be critical for team performance (Mehra et al., 2006; Gerpott et al., 2019). The leadership role represents a boundary inside the team and performance is improved when individuals inside the boundary (i.e., team members who are primarily responsible for coordination) are connected and individuals outside the boundary (i.e., team members who represent different perspectives) are disconnected.[3] The network imagery is identical to the network described by Reagans and Zuckerman. Our results are consistent with broader set of findings on teams in the field and have the added benefit of being causal.

Our research has implications for scholars with an interest in network-based benefits defined more broadly. Our contrast between CLIQUE and WHEEL teams reflects an on-going debate between two network forms of social capital. The debate focuses on the benefits provided by network closure (a CLIQUE network) versus the benefits provided by a structural hole or brokerage network (a network organized around a central individual). The two network forms are usually viewed in opposition (Burt 2000 provides a systematic review). In a very general

[3] An underappreciated feature of a CLIQUE network is flexibility. With experience working together, team members could discover how they should be organized. For example, if a team is performing a basic task it could decide to turn a decentralized network into a centralized network (Guetzkow and Simon, 1955).

way, the benefits provided by a structural hole network come at the expense of the benefits provided by a closed network. A network connection can either contribute to more structural holes in a network or more closure but not both. While it can be difficult for an individual to develop a network that realizes the benefits of both networks, as we have noted above, the tension is alleviated in groups and teams. The tension can be alleviated in larger collectives in general, including organizations (Lazer and Friedman, 2007; Fang et al., 2010) and even markets (Burt, 2000: 392-398). Our findings are consistent with a more integrative view of brokerage and closure (Burt, 2005).

The superiority of the CB network is more than an integration of brokerage and closure. The disconnects between the peripheral members allow each one to develop relatively distinct interpretations of the symbols. To do well, each broker has to interpret the ambiguous information he or she receives from each peripheral team member, which is difficult. Making sense of the ambiguous information is aided by the fact that the brokers are connected to the same people. The brokers receive the same information and the tie between them allows them to coordinate their interpretations and reach a mutual understanding of what they are hearing. The ties from the brokers to each peripheral member represent a “wide bridge” that facilitates the interpretation of complex information. Indeed, the wide bridge in our CB network is a “Simmelian” tie (Krackhardt, 1999). We are certainly not the first researchers to emphasize the importance of Simmelian ties for knowledge transfer. Tortoriello and Krackhardt (2010) describe how Simmelian ties help individuals, in their case knowledge brokers, overcome their outsider status in different technological domains, allowing them to engage more fully in the transfer of knowledge across domains, ultimately enhancing their capacity for creativity and innovation. Our results illustrate the importance of Simmelian ties for interpreting and integrating ideas.

The findings have important implications for how we think about network advantage in general.

An emerging body of research defines brokering advantages in terms of synthesizing complex knowledge and information (Burt, 2020). For example, individuals who bridge structural holes derive a greater capacity for the transfer of more complex knowledge and information (Reagans and McEvily, 2003; Tortoriello, Reagans, and McEvily, 2012). We tend to view brokerage as an individual activity. Perhaps, we should not. During the initial stages of a market, for example, brokers can legitimate each other's efforts (Burt, 2005:chap. 5). Our findings suggest the same brokers could benefit from coordinating their problem-solving efforts. Wide bridges can give brokers an advantage at interpreting and integrating knowledge and information (Ter Wal et al., 2020). Our results provide insight into how individuals reach consensus when they are working in an area where they cannot simply depend on their prior knowledge and expertise, but instead must figure out what to do together.

References

Aral, S. and Van Alstyne, M., 2011. The diversity-bandwidth trade-off. *American journal of sociology*, *117*(1), pp.90-171.

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

Balkundi, P. and Harrison, D.A., 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), pp.49-68.

Barber, S.J., Harris, C.B. 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), p.559.

Bavelas, A. 1950. "Communication Patterns in Task-Oriented Groups." The Journal of the Acoustical Society of America 22 (6): 725–30.

Bayram, A.B. and Ta, V.P., 2019. Diplomatic chameleons: Language style matching and agreement in international diplomatic negotiations. *Negotiation and Conflict Management Research*, *12*(1), pp.23-40.

Burgess, R.L., 1968. Communication networks: An experimental reevaluation. *Journal of Experimental Social Psychology*, *4*(3), pp.324-337.

Burt, R.S., 1987. Social contagion and innovation: Cohesion versus structural equivalence. *American journal of Sociology*, *92*(6), pp.1287-1335.

Burt, R.S., 2000. The Network Structure of Social Capital. Research in Organizational Behavior. Vol. 22.

Burt, R.S., 2005. Brokerage & Closure. New York: Oxford University Press.

Burt, R.S., 2020. Structural Holes Capstone, Cautions, and Enthusiasms. In *Personal Networks: Classic Readings and New Directions*, edited by Brea L. Perry, Bernice Pescosolido, Mario L. Small, and Ned Smith.

Burt, R. S., Reagans, R. E. 2020. "Team talk: The network origins of jargon." Working paper, The University of Chicago.

Burt, R. S., Reagans, R. E. and Volvovsky H.C.,. 2020. "Network brokerage and the perception of leadership." *Social Networks*, forthcoming.

Christie, L.S., Luce, R.D. and Macy, J., 1952. Communication and learning in task-oriented groups.

Clark, H.H. and Wilkes-Gibbs., D. 1986. “Referring as a Collaborative Process.” Cognition 22 (1): 1–39.

Cross, R., Rebele, R. and Grant, A., 2016. Collaborative overload. *Harvard Business Review*, *94*(1), p.16.

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

DellaPosta, D., Shi, Y. and Macy, M., 2015. Why do liberals drink lattes?. *American Journal of Sociology*, *120*(5), pp.1473-1511.

Diehl, M. and Stroebe, W., 1987. Productivity loss in brainstorming groups: Toward the solution of a riddle. *Journal of personality and social psychology*, *53*(3), p.497.

Fang, C., Lee, J. and Schilling, M.A., 2010. Balancing exploration and exploitation through structural design: The isolation of subgroups and organizational learning. *Organization Science*, *21*(3), pp.625-642.

Faraj, S. and Sproull, L., 2000. Coordinating expertise in software development teams. *Management science*, *46*(12), pp.1554-1568.

Fraley, C. and Raftery, A.E., 2002. Model-based clustering, discriminant analysis, and density estimation. *Journal of the American statistical Association*, *97*(458), pp.611-631.

Gonzales, A.L., Hancock, J.T. and Pennebaker, J.W., 2010. Language style matching as a predictor of social dynamics in small groups. *Communication Research*, *37*(1), pp.3-19.

Gerpott, F.H., Lehmann-Willenbrock, N., Voelpel, S.C. and Van Vugt, M., 2019. It’s not just what is said, but when it’s said: A temporal account of verbal behaviors and emergent leadership in self-managed teams. *Academy of Management Journal*, *62*(3), pp.717-738.

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

Hoffer Gittell, J., 2002. Coordinating mechanisms in care provider groups: Relational coordination as a mediator and input uncertainty as a moderator of performance effects. *Management science*, *48*(11), pp.1408-1426.

Huckman, R.S., Staats, B.R. and Upton, D.M., 2009. Team familiarity, role experience, and performance: Evidence from Indian software services. *Management science*, *55*(1), pp.85-100.

Hummon, N.P., Doreian, P. and Freeman, L.C., 1990. Analyzing the structure of the centrality-productivity literature created between 1948 and 1979. *Knowledge*, *11*(4), pp.459-480.

Katz, N., Lazer, D., Arrow, H. and Contractor, N., 2004. Network theory and small groups. *Small group research*, *35*(3), pp.307-332.

Krackhardt, D., 1999. The ties that torture: Simmelian tie analysis in organizations. SB Andrews, D. Knoke, eds. Networks in and Around Organizations. Research in Sociology of Organizations, Vol. 16.

Lazer, D. and Friedman, A., 2007. The network structure of exploration and exploitation. *Administrative science quarterly*, *52*(4), pp.667-694.

Le, Q. and Mikolov, T., 2014, January. Distributed representations of sentences and documents. In *International conference on machine learning* (pp. 1188-1196).

Leavitt, H.J., 1949. Some Effects of Certain Communication Patterns upon Group Performance. Doctoral Dissertation, Massachusetts Institute of Technology, Boston, MA.

Leavitt, H.J., 1951. Some effects of certain communication patterns on group performance. *The Journal of Abnormal and Social Psychology*, *46*(1), pp.38-50.

Levy, O. and Goldberg, Y., 2014. Neural word embedding as implicit matrix factorization. In *Advances in neural information processing systems* (pp. 2177-2185).

Lix, K., Goldberg, A., Srivastava, S. and Valentine, M.A., 2020. Timing Differences: Discursive Diversity and Team Performance. *SocArXiv. June*, *12*.

March, J.G. and Simon, H.A., 1958. Organizations John Wiley & Sons. *New York*.

Mehra, A., Dixon, A.L., Brass, D.J. and Robertson, B., 2006. The social network ties of group leaders: Implications for group performance and leader reputation. *Organization science*, *17*(1), pp.64-79.

Mikolov, T., Chen, K., Corrado, G. and Dean, J., 2013. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Mora-Cantallops, M. and Sicilia, M.Á., 2019. Team efficiency and network structure: The case of professional League of Legends. *Social Networks*, *58*, pp.105-115.

Mukherjee, S., 2016. Leadership network and team performance in interactive contests. *Social Networks*, *47*, pp.85-92.

Mundlak, Y., 1978. On the pooling of time series and cross section data. *Econometrica: journal of the Econometric Society*, pp.69-85.

Putman, V.L. and Paulus, P.B., 2009. Brainstorming, brainstorming rules and decision making. *The Journal of creative behavior*, *43*(1), pp.29-40.

Rajaram, S. and Pereira-Pasarin, L.P., 2010. Collaborative memory: Cognitive research and theory. *Perspectives on psychological science*, *5*(6), pp.649-663.

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), pp.869-881.

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

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

Reagans, R. and Zuckerman, E.W., 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), pp.101-133.

Rogge, G.O., 1953. Personality factors and their influence on group behavior: a questionnaire study.

Saint-Charles, J. and P. Mongeau. 2018. Social influence and discourse similarity networks in workgroups. *Social Networks*, *52*, pp.228-237.

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), pp.7361-7366.

Shaw, M.E., 1954. Some effects of unequal distribution of information upon group performance in various communication nets. *The journal of abnormal and social psychology*, *49*(4p1), p.547.

Shaw, M.E., 1964. Communication networks. In *Advances in experimental social psychology* (Vol. 1, pp. 111-147). Academic Press.

Shore, J., Bernstein, E. and Jang, A.J., 2020. Network Centralization and Collective Adaptability to a Shifting Environment. *Available at SSRN 3664022*.

Smith, S.L., 1950. Communication pattern and the adaptability of task-oriented groups: an experimental study. *Group Networks Laboratory, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge*.

Srivastava, S.B., Goldberg, A., Manian, V.G. and Potts, C., 2018. Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations. *Management Science*, *64*(3), pp.1348-1364.

Tausczik, Y.R. and Pennebaker, J.W., 2013, April. Improving teamwork using real-time language feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 459-468).

Taylor, D.W., Berry, P.C. and Block, C.H., 1958. Does group participation when

using brainstorming facilitate or inhibit creative thinking?. *Administrative Science Quarterly*, pp.23-47.

Taylor, P.J. and Thomas, S., 2008. Linguistic style matching and negotiation outcome. *Negotiation and Conflict Management Research*, *1*(3), pp.263-281.

Ter Wal, A.L., Criscuolo, P., McEvily, B. and Salter, A., 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), pp.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), pp.1024-1039.

Weber, R.A. and Camerer, C.F., 2003. Cultural conflict and merger failure: An experimental approach. *Management science*, *49*(4), pp.400-415.

Wilkes-Gibbs, D. and Clark, H.H., 1992. Coordinating beliefs in conversation. *Journal of memory and language*, *31*(2), pp.183-194.

Wooldridge, J.M., 2010. *Econometric analysis of cross section and panel data*. MIT press.

Figure 1: Centralization, Language Similarity, and Performance

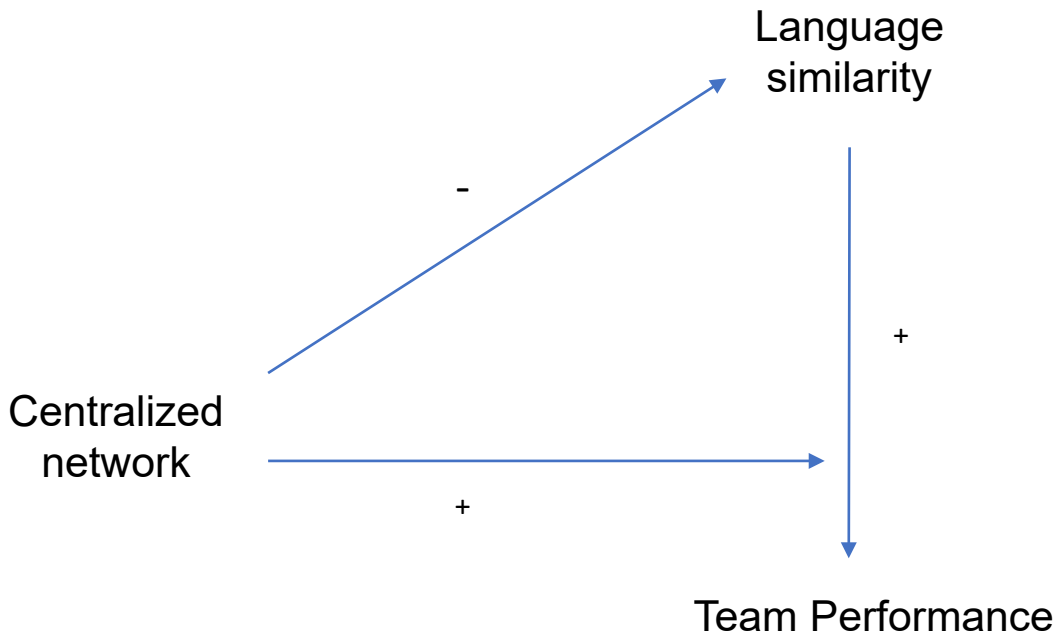
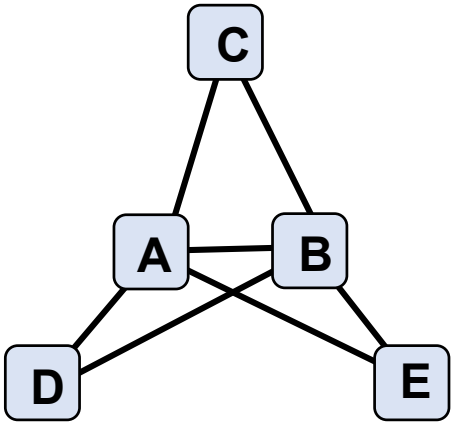
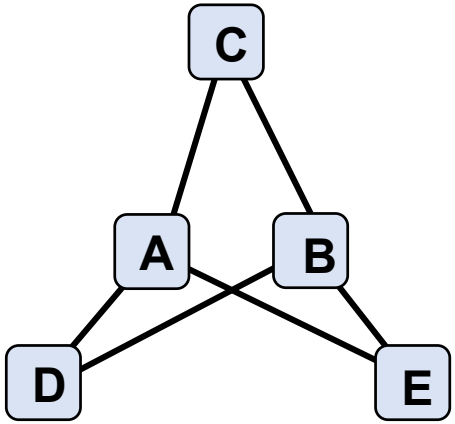


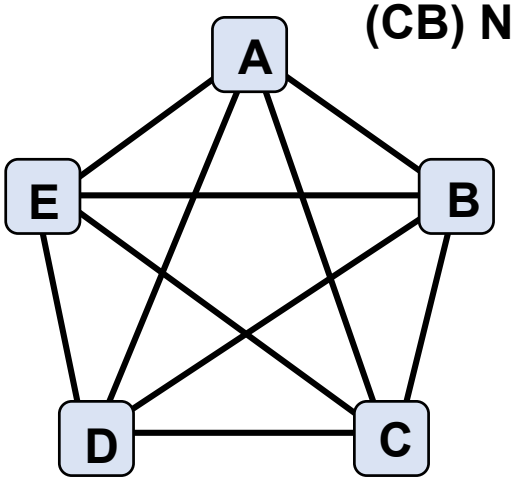
Figure 2. Four Team Networks.



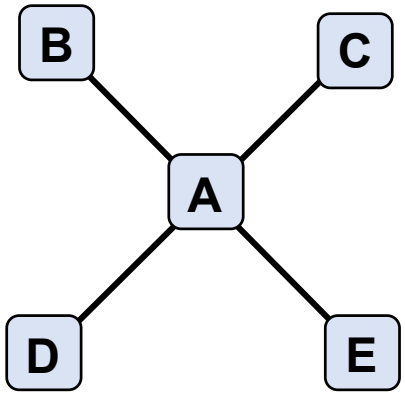
Connected
Brokers
(CB) Network



Disconnected
Brokers
(DB) Network

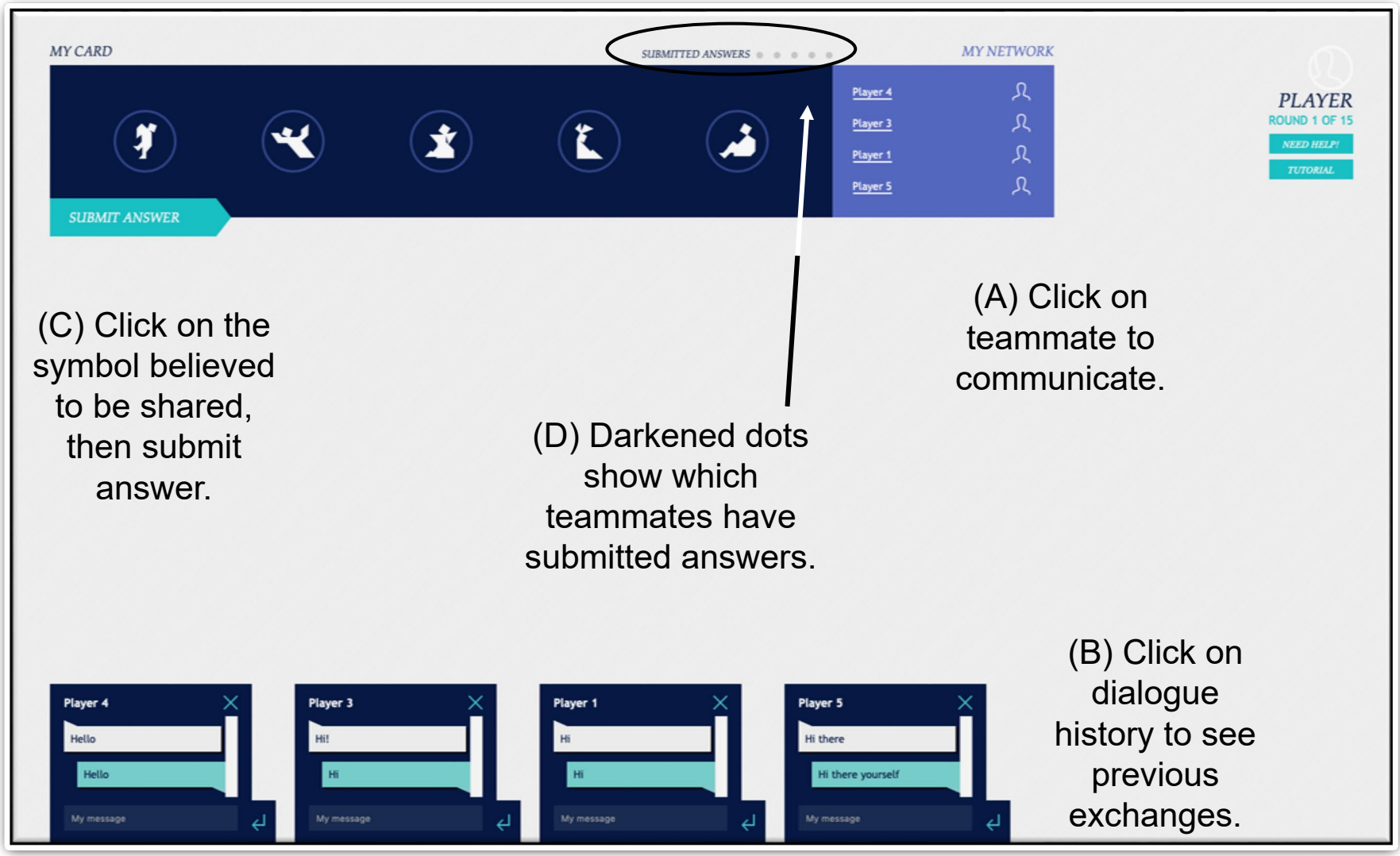


CLIQUE
Network



WHEEL
Network

Figure 3. Example Game Screen.



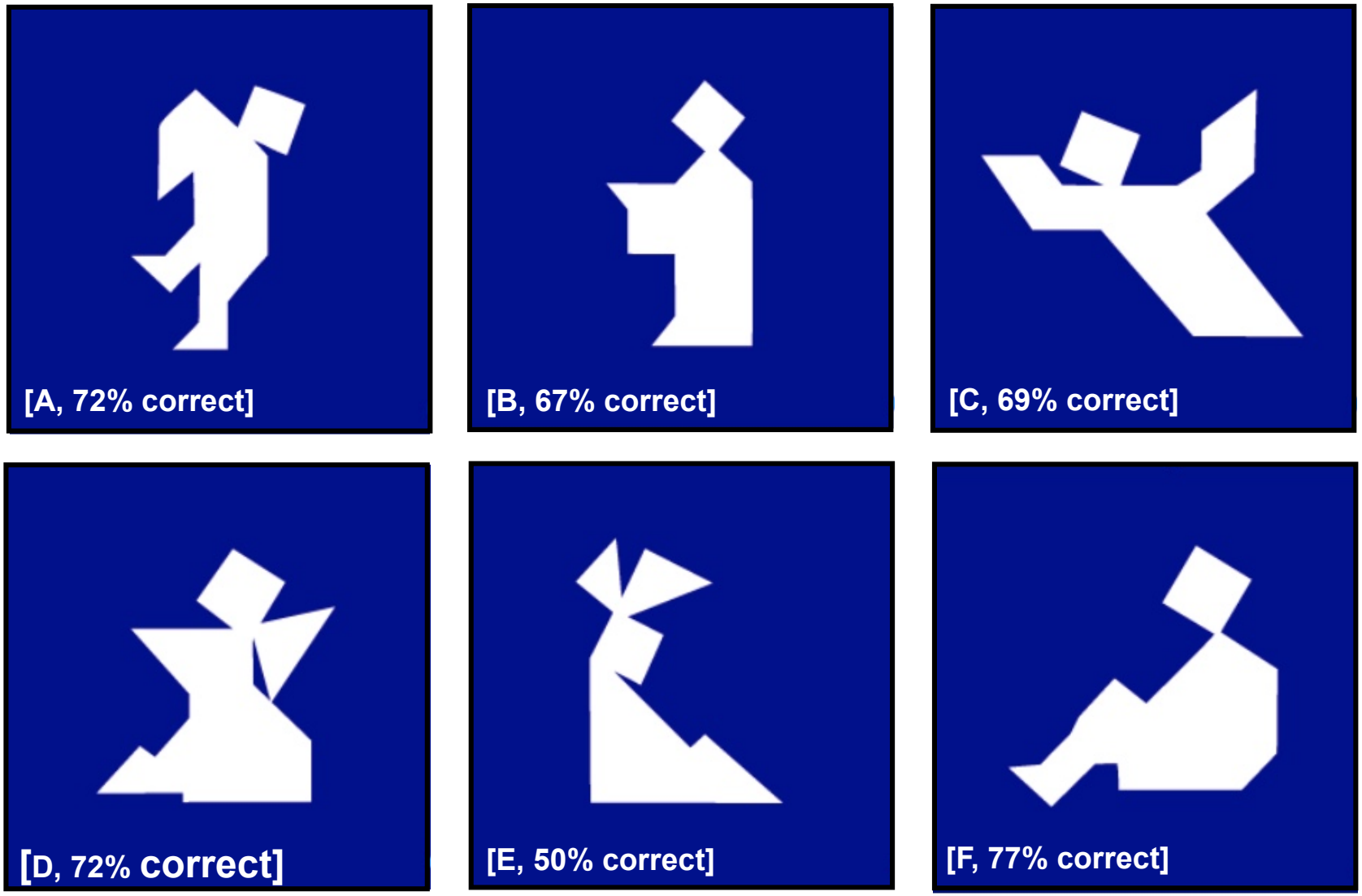
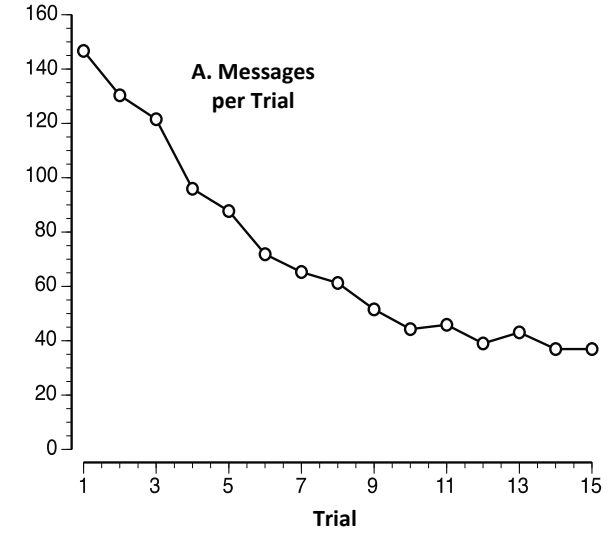
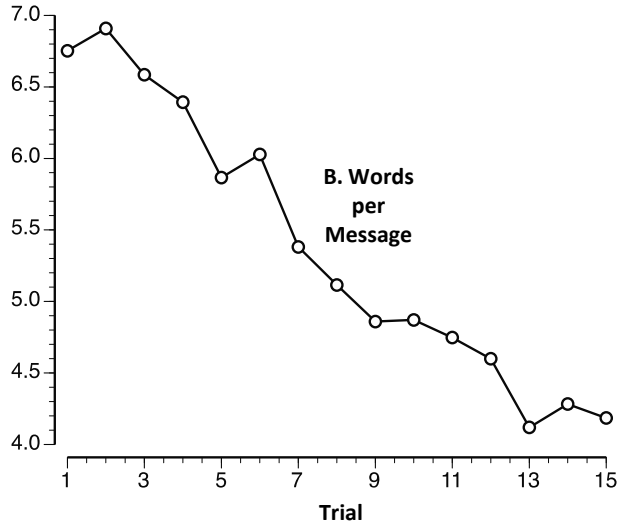


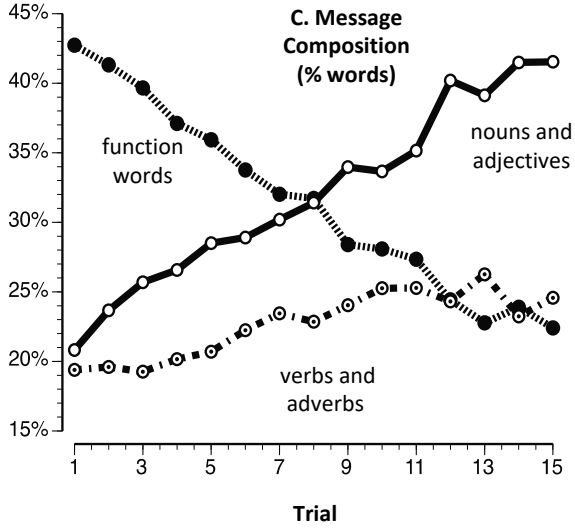
Figure 4. Subject's Hand Is Five of These Tangram Figures.

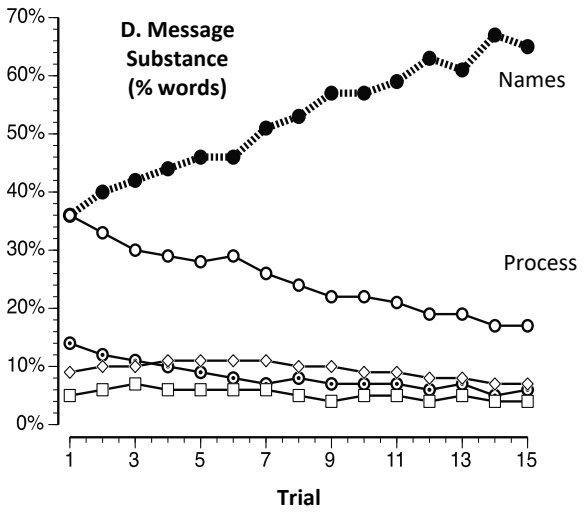
NOTE — Identification in brackets did not appear on game screen (see Figure 3).

Figure 5. Team Language Becomes More Efficient, Specialized, and Similar Across Trials









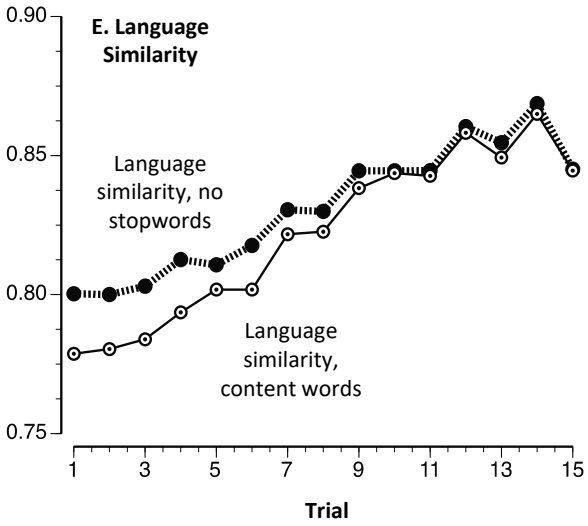


Table 1: Word Clusters

	Process	Name	Feature	Action	Misc.
Representative words and phrases	1. **think**: “i think g6 and g9 have the priest as well” 2. **symbol**: “which other symbols do you have?” 3. **everyone**: “seems like one leg up is shared by everyone”	1. **man**: “scroll man, waving arms man, offset triangles, rabbit, man sitting on ground” 2. **bunny**: “separate legts, dimploma, warrior, bunny laying down” 3. **priest**: “im confused my the priest and the waiter”	1. **right**: “two legs to the left are attached to a trapezoid body and there's a tilted square to the right” 2. **triangle**: “3. square head has two triangles coming out of it, body is a right triangle” 3. **bottom**: “body is triangle w/ vertical edge on left side and horizontal on bottom”	1. **facing**: “do you have a native cheif facing left” 2. **kneeling**: “Ghost, child kneeling behind rock with raised hands, rabbit looking backwards, one legged man laying down, thing with 2 legs (dancing?)” 3. **guess**: “ I guess the rabbit ears could look like a guy kneeling with arms outstretched ”	1. **yup**: “yup!” 2. **k**: “k” 3. **ya**: “ya” or “ya $$$”
Number of words in cluster	471	415	260	551	566
Frequency of use	51643	290720	21619	39281	10348

Table 2: Descriptive 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) Language similarity	.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) MIT	.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) CLIQUE	.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	.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

Table 3: Team Network, Language Similarity, and Team Accuracy

	Column I	Column II	Column III Equation 1	Column III Equation 1	Column IV Equation 2	Column IV Equation 2
	Language Similarity	Language Similarity	Language Similarity	Team Accuracy	Language Similarity	Team Accuracy
Symbol1	-0.002	-0.001	-0.001	-0.328*	-0.001	-0.303*
	(-0.44)	(-0.31)	(-0.30)	(-2.40)	(-0.30)	(-2.20)
Symbol2	-0.002	-0.002	-0.002	-0.137	-0.002	-0.135
	(-0.49)	(-0.54)	(-0.53)	(-1.00)	(-0.53)	(-0.99)
Symbol3	-0.011*	-0.011*	-0.011*	-1.010***	-0.011*	-1.015***
	(-2.41)	(-2.43)	(-2.42)	(-7.43)	(-2.42)	(-7.47)
Symbol4	0.001	0.001	0.001	-0.187	0.001	-0.149
	(0.20)	(0.25)	(0.29)	(-1.38)	(0.29)	(-1.09)
Symbol5	-0.005	-0.005	-0.005	0.331*	-0.005	0.313*
	(-1.08)	(-1.07)	(-1.05)	(2.23)	(-1.04)	(2.11)
MIT	0.003	0.003	0.003	0.134	0.003	0.178*
	(0.40)	(0.42)	(0.40)	(1.53)	(0.40)	(2.02)
CB network	0.029**	-0.007	-0.007	0.013	-0.007	-0.467
	(3.14)	(-0.55)	(-0.53)	(0.10)	(-0.53)	(-0.19)
DB network	0.006	-0.014	-0.014	0.041	-0.014	2.312
	(0.72)	(-1.21)	(-1.20)	(0.39)	(-1.20)	(1.02)
CLIQUE	0.019*	-0.006	-0.006	-0.471***	-0.006	8.063***
	(2.13)	(-0.53)	(-0.54)	(-4.20)	(-0.54)	(3.83)
Log trial	0.023***	0.012***	0.011***	0.793***	0.011***	0.809***
	(13.80)	(3.52)	(3.47)	(15.64)	(3.47)	(15.74)
CB X Log trial		0.020***	0.020***		0.020***	
		(4.33)	(4.29)		(4.29)	
DB X Log trial		0.011*	0.011*		0.011*	
		(2.54)	(2.53)		(2.53)	
CLIQUE X Log trial		0.015**	0.015**		0.015*	
		(3.26)	(3.29)		(3.28)	
Language similarity (LS)				9.187***		13.170***
				(9.37)		(6.52)
CB X LS						0.462
						(0.15)
DB X LS						-2.858
						(-1.01)
CLIQUE X LS						-10.490***
						(-4.03)
Tenth trial	0.009	0.009	0.009	0.776***	0.009	0.841***
	(1.69)	(1.71)	(1.70)	(3.57)	(1.70)	(3.78)
Team collapse	-0.016*	-0.017*	-0.016*	-2.506***	-0.016*	-2.525***
	(-2.25)	(-2.34)	(-2.20)	(-12.62)	(-2.20)	(-12.55)
Constant	0.774***	0.794***	0.794***	-7.251***	0.794***	-10.490***
	(81.97)	(74.77)	(74.36)	(-9.27)	(74.37)	(-6.48)

t statistics in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$