Forthcoming in *Organization Science*

PHANTOM NETWORKS:

Window on Agency in Social Mechanisms

June, 2026

Ronald S. Burt
Bocconi University, Milan, IT 20136
rburt@uchicago.edu
ORCID: 0000-0002-7947-2101

Ray E. Reagans
MIT, Boston, MA 02142
rreagans@mit.edu
ORCID: 0000-0001-8706-3124

Sonja Opper
Bocconi University, Milan, IT 20136
sonja.opper@unibocconi.it
ORCID: 0000-0001-9181-8323

Keywords: Agency, Experimental Design, Motivation, Leadership, Social Mechanism, Social Networks,

ACKNOWLEDGEMENT: This work was made possible by financial support from the University of Chicago Booth School of Business for data-processing software, and financial support from MIT Sloan for transitioning our earlier experiment to the more flexible Empirica platform, and funding the experiment trials on which we report here. We are grateful to Ross Stolzenberg for his helpful comments during the analysis, and to Matthew Bothner, Damon Centola, Alessandro Iorio, Martin Kilduff, Ajay Mehra, Eric Quintane, Michelle Rogan, Stefano Tasselli, and Tom Valente for their comments on the manuscript. The text has also benefitted from discussion in symposia at the 2023 and 2024 meetings of the Academy of Management, the 2024 meetings of the Social Network Society, and a seminar at ESMT.

PHANTOM NETWORKS:

Window on Agency in Social Mechanisms

ABSTRACT

Our goal is to offer proof of concept for the idea that behavioral networks emerging with and beyond assigned treatment networks in experiments can provide a valuable window on agency. We call these often unobserved behavioral structures phantom networks. Most evidence of agency is entangled with audience — what people recall, how others attribute, or who is seen to act. We strip that away to observe agency directly in the ties participants build within and beyond what they are assigned. Their deviation from assignment is not error but structure-in-action, a behavioral trace of agency. We offer three exhibits of that "behavioral deviation" revealing agency in gendered brokerage, rival brokerage, and structural reproduction. Across the three exhibits, behavioral deviation shows how a team works (sometimes in surprising ways), not how well the team works. Teams perform as expected in theory, but the behavior responsible for their performance is made more apparent by watching behavioral deviation. We close with a summary proof of concept and discussion of next steps for particularly promising future work.

At the heart of all research is the question “How does it work?” In social science, the question often manifests as a concern with social mechanisms, the processes by which what we see comes to be. Particularly prominent in this concern is Coleman’s recommendation that we explain correlations between macro conditions in terms of micro processes among actors using resources (popularly known as the “Coleman boat,” Coleman, 1990: 8-11, 702). Hedstrom and Swedberg (1998) offer energetic exploration of the “boat” on diverse topics. Central in such exploration is the concept of agency — the capacity of people to act so as to reproduce, capitalize upon, or alter the situation in which action occurs. Moving closer to this paper, agency is especially important in research on social networks, wherein outcomes depend not only on structure, but on what actors do within structure (Tasselli and Kilduff 2021). Vissa (2012), for example, shows that network outcomes depend on available ties and how actors mobilize those ties. The centrality of agency, however, is matched by the difficulty in observing it (Emirbayer and Mische, 1998; Sewell, 1992). The difficult question is not whether agency matters, but how to observe it. In most work, agency and mechanism are inferred from roles, recall, or attribution — what people say they did, or what audiences say people did.

As a lens on agency in social mechanisms, we use experiments. Specifically we use experiments involving interpersonal behavior, that is to say, network experiments. In a network experiment, two or more people engage in some form of coordination that is subject to limits on who can interact with whom. The allowed interaction defines a treatment network. Alternative patterns of interaction can be studied for their consequences by randomly assigning people to alternative treatment networks, then looking for different results from people assigned to different treatments.

Network experiments contain more than treatment networks. Participants also bring personal networks into the experiment and create new behavioral networks during the experiment. The result is a three-network lens on behavior: the treatment network, the personal network, and the behavioral network. Standard analyses focus on the first. In the spirit of Webb et al.’s (1966) “unobtrusive measures,” we focus on the latter two, often ignored structures that coexist with treatment in every network experiment. The personal network describes a participant’s interpersonal behavior outside the experiment. The behavioral network is the network of interpersonal behavior that emerges within the experiment. Respectively external and internal behavior, the two networks are phantoms in standard research practice (cf. Greenberg, 2021:1178). By the term “behavioral deviation” we refer to the extent to which participants’ behavioral networks depart from the structure defined by treatment.

We here measure behavioral deviation by the constraint on a person in the behavioral network minus constraint defined by treatment. In other words, is a person’s behavioral network less closed than the network to which he or she is assigned by treatment? Negative values indicate agentic behavior that allowed someone to wiggle free from treatment; larger negative values indicating larger decreases in constraint.

In the language of experiment design, behavioral deviation is a form of bias or noncompliance. As such, it is familiar across the social sciences: Students often get different benefit from education experiments depending on how they engage program content (Orosz et al. 2021). Police officers instructed in an experiment to respond in specific ways to specific circumstances, sometimes adapt their response to what they believe is appropriate to the situation (Angrist 2006). Medical treatment can vary with facilities available in the nearest hospital during an emergency (McClellan et al. 1994), and management interventions can involve participants absent or uninvited (Carnabuci and Quintane 2023). Reviews are readily available on kinds of people and situations prone to noncompliance (e.g., Ali et al. 2021; DiMatteo 2004; Jancsics et al. 2023; Zolnierek and DiMatteo 2009).

It follows from familiarity that strategies exist to make inferences despite noncompliance (e.g., Little and Yau 1998; Sagarin et al. 2014). The most obvious is to ignore it, which can be expected to obscure treatment effects, providing a lower-bound estimate of treatment effect.[1] An opposite extreme is to eliminate noncompliant participants. People likely to be noncompliant can be removed from the data using attention checks, or knowledge of participant attributes strongly associated with noncompliance, or post-hoc removal of participants discovered to

[1]Ignoring noncompliance is characteristic of early network experiments, is often found in current work, and is not without justification in policy research. The treatment effect is biased toward zero, but if a certain amount of noncompliance is to be expected in the target population for a policy, then the treatment effect estimated ignorant of noncompliance is a realistic estimate of the policy effect (see discussions of “intention to treat” analysis).

have been noncompliant. This strategy can provide an upper-bound estimate of treatment effect when the goal is to study how an effect works when it works. The complication is that this strategy can generate selection bias when specific categories of people are prone to noncompliance. Between the two strategy extremes, statistical strategies include analysis of covariance models that hold known compliance factors constant when estimating network effects, or two-stage models that use treatment to define instrumental variables to estimate effects from indices defined by behavioral networks (Carnabuci and Quintane 2023; Reagans and Burt 2026).

Here we propose phantom networks as a research opportunity. The gap between assigned and behavioral structure is a window on agency. The two behavioral networks constitute or shape how participants experience treatment. Treatment networks define which channels of interaction are available. They do not define how attention is distributed across allowed channels. And if treatment is to resemble socializing outside the experiment, treatments should not impose such restrictions. Free to choose between allowed contacts, a preference for communication with certain teammates is expected for a variety of reasons: Their language, written or verbal, is more familiar to you. They display familiarity with a topic that interests you. They are more courteous, more earnest. Felt homophily can emerge from many origins. Whatever the reason, certain teammates can become more appealing targets of communication. Other connections defined by treatment can atrophy for want of attention. These behavioral deviations from treatment are valuable data. They are a trace on individual preference, discomfort, and initiative. They show when people accept a treatment structure, and when they bend it. They reveal structure emergent.

Five sections of text follow. We first sketch replication of an iconic coordination experiment on team leadership from which we draw illustration. The subsequent three sections are exhibits of phantom networks used to explore agency. We consider gender, rivalry, and today's network predicting tomorrow's. Each section consists of a question about agency, evidence on the question, and a demonstration that controlling for the observed agency does not affect team performance. That last point is to be noted. We find that knowledge of how a team works is separate from how well it works.

The exhibits reveal a deeper understanding of social mechanisms in team performance. In gender, we find that informal leadership emerges even when gender is invisible — not because men are assigned broker roles, but because they assert it. In rivalry, we find that structure alone is not enough. Rivalry appears only when network structure renders similarity visible and actionable. And for structural reproduction, we find that prior brokerage experience does not predict brokerage behavior; quite the opposite. Participants from closed networks are more likely to emerge from among teammates as brokers when they can get away with it.

Last, we close with summary proof of concept for behavioral deviation as a window on agency, and discuss promising lines for future work with behavioral deviation. The action implication for people who run network experiments is a recommendation to make use of data on the behavioral network a participant brings into the experiment, and the behavioral network that emerges during the experiment. The implication for people interested more generally in agency and social mechanisms is a recommendation to ask whether colleagues advocating a mechanism have evidence of the mechanism operating or are merely — as is usually the case — making inference from post-hoc opinion about what happened.

At this point, you might want to jump ahead to the exhibits. We have written in the spirit of an *Organization Science* Perspective discussion, which aims at a diverse audience. We find the exhibits exciting for the potential they display, but transparency is important too. Details of the experiment and measurement of the three networks are provided in the next section for readers interested in knowing how we obtained the evidence reported in the exhibits. Replication work will certainly depend on both experiment and exhibits.

EMPIRICAL SETTING: EXPERIMENT AND DATA

We draw illustrative data from a renovation of the iconic Bavelas-Smith-Leavitt team coordination experiment (Leavitt 1951; Hummon et al. 1990; Leavitt 1996; Freeman 2004:66-71). The familiar design is broadly relevant to structure as it affects coordination and achievement — viz., the related network models of centrality (Freeman 1979), power/status (Coleman 1990:667ff.; Mizruchi et al. 1986; Podolny 1993, 2005), and structural holes (Burt 1992, 2005).

The experiment involves a team of five people, each of whom is given a “hand” of five symbols drawn from a population of six symbols. The task is for teammates to communicate with one another to discover which symbol they have in common. The task would be simple with symbols easily communicated. The original experiment used familiar symbols (circle, square, triangle, etc.). Burt et al. (2021:37-38), drawing on the “cloudy marbles” variant of the early experiments at MIT (Christie et al. 1952), make the task more complex by using abstract symbols, so called “tangrams” (which we also use in our replication, see top of Figure 1 from Reagans et al., 2026:4). To make exploration more practical, we re-created the experiment on the Empirica platform (https://empirica.ly), drawing participants from the Prolific population (Palan and Schitter 2018; https://www.prolific.co).

——— Figure 1 About Here ———

Teams: Adapted from Burt et al. (2021:39), Figure 2 displays the treatment networks and the seven positions to which participants are assigned at random. Squares represent participants. Lines connect pairs of teammates allowed to communicate with one another. Overall, we aimed for 20-25 teams in each network condition.

The primary contrast is between the WHEEL network, in which one person has monopoly control of communication between teammates (position 1, brokering communication across 6 structural holes between teammates), versus the CLIQUE network, in which every teammate has access to one another, so no one is a broker. The two mixed networks are a pair of disconnected brokers connected to the same contacts (position 2), and a pair of connected brokers who have the same contacts (position 4, alternatively the treatment is a pair of overlapping cliques). We will refer to the connected-brokers treatment as the CB network and the disconnected-brokers treatment as the DB network.

——— Figure 2 About Here ———

The task: Teammates have to coordinate on words with which they can identify the tangrams. From Reagans et al. (2026:7), Figure 3 shows the participant-machine interface. Teammates do not see or hear one another. Communication is by text message. For easy reference, teammates identify one another by color: Black, Blue, Green, Purple, Red, White, and Yellow. A participant’s assigned color and position remain constant through the experiment. Displayed across the bottom of the screen are teammates in contact with the participant. Click on a teammate to send a message to the teammate (point A in Figure 3). The example interface in Figure 3 displays four teammates, so it is for a participant assigned to the hub position in a WHEEL network, or one of the connected brokers in a CB network, or any position in a CLIQUE network (respectively positions 1, 4, and 5 in Figure 2). At the opposite extreme, a participant assigned to position 7 has only one teammate displayed.

——— Figure 3 About Here ———

The five tangrams in the participant's current hand are displayed across the top of the screen. To submit a guess about the symbol all teammates share, click on the symbol, which puts a circle around it as in Figure 3, then click on the submit button (point B in Figure 3). Across the top of the screen are bits of information to encourage participants to move along (point C in Figure 3). When the time for this trial runs out (1/15 of the one-hour total experiment time), a bell sounds to indicate that further time on this trial is coming from the budget for subsequent trials.

Participants: Participants are screened volunteers in the US. The screening runs as follows: Interested Prolific participants sign up for the study lasting up to an hour with compensation of eight dollars plus performance bonuses (one dollar per teammate for each correct team solution). In an introductory session, they answer 10 questions to assess their English language skills. Those who pass the language screen are given a tutorial on doing the experiment task, then asked comprehension questions. Those who pass the comprehension screen are presented with an example tangram and asked to write how they would describe the tangram to another person, then they are asked the pre-experiment name generator and name interpreters to describe their core personal network. At this point, the person is invited to participate in the experiment. There are no quotas for the usual sampling strata of gender, race, education, or age.

The final assembly of 465 participants are 51% male, primarily high-school (40%) or college (40%) graduates, varying in age from 18 to 73 around a mean of 38 years. We collected data in batches of 8-12 teams during February and March, 2023. A trial begins when each teammate receives his or her "hand," and continues until all five teammates have submitted their guess of the shared symbol. The experiment consists of 15 trials.

Personal Network Pre-Experiment

Figure 1 displays the text of the name generator we used to measure participants' personal networks pre-experiment. Patterned on the widely-used General Social

Survey name generator (Burt 1984; Marsden 1987; Marsden et al. 2021), the question asks for people with whom a participant has most often discussed important matters in the last six months. The participant is encouraged to use initials or nicknames, and told that the names will not be recorded.

The names are used immediately in two name interpreter questions, then replaced with sequential numbers in the recorded data. The name interpreters ask how often participants spoke with each named discussion partner ("daily," "weekly," "less often"), and how often each pair of discussion partners spoke with each other. These data are sufficient to compute a network constraint score for each participant (Burt 1992: 64).[2] Constraint decreases toward zero in larger networks (degree or network size), and increases with the strength of connections between contacts (density), or concentration of relations in a subset of contacts (hierarchy or centralization). The higher the network constraint, the more a participant's interpersonal behavior outside the experiment is within a closed network. We multiply constraint by 100 to speak in terms of points of constraint, and Winsorize constraint in small dense networks to 100.

Example networks of high and low constraint networks are displayed in Figure 1. Squares are people. Lines are connections. Thicker lines indicate more frequent discussion.[3] The sociogram to the left shows a person whose discussion partners rarely talk to one another. The sociogram to the right shows a person whose discussion partners mostly talk with one another every day. With constraint scores grouped into five-point intervals, the histogram in Figure 1 shows a roughly-normal

---

[2]The network constraint on a person i is the sum of constraint from each of *i*'s contacts ($C_i = \Sigma_j c_{ij} = \Sigma j (p_{ij} + \Sigma_k p_{ik}p_{kj})^2$, $I \neq k \neq j$, where p_{ij} is the proportion of *i*'s network communication that is with contact *j* ($\Sigma_j p_{ij} = 1.0$). Person *i* is constrained in his or her relationship with teammate *j*, c_{ij}, to the extent that *i* cannot avoid the teammate — because a large portion of *i*'s communication is with *j* (high p_{ij}), and the teammate is strongly connected with everyone else in the team ($\Sigma_k p_{ik}p_{kj}$, $i \neq k \neq j$).

[3]Participants had three response options to describe contact frequency: daily, weekly, or less often. We computed network constraint from interval and scaled scores. The interval scores are 3, 2, 1 respectively for the three levels of frequency. We used balance theory to scale the intervals (Burt 2010:290-293), which resulted in the following scores for daily, weekly, and less often: 1.00, .36, and .06 in relations with discussion partners, and .96, .45, and .00 in relations between discussion partners. Since the interval scores do not discount "less often" as much as the scaled scores, network constraint computed from the interval scores is slightly higher (60.97 mean versus 55.64 respectively). However, correlations with assigned and behavioral networks are similar with either scaling. Constraint in personal networks defined by interval scoring is correlated with constraint in the assigned and behavioral network during the experiment (*r* = .09 and .08, respectively). The same correlations are *r* = .10 and .07 when constraint in personal networks is defined by scaled scores. Given the small difference, we use the scaled scores because they are a bit more informed by the way participants used the response categories.

distribution of constraint varying from 19 to 99 points around an average of 55.64 (*SD* = 16.61). These are not extensive data on pre-experiment networks, but they are sufficient to distinguish participants by the extent to which they enter the experiment accustomed to interpersonal behavior in closed networks, with its normative prescription for proper behavior and preservation of the network status quo (Burt 2002, 2005; Martin and Yeung 2006; Martin et al. 2022).

Treatment Network

We use network constraint to measure the extent to which a participant's position in a treatment network provides brokerage opportunities. Access to a structural hole occurs when a participant is connected to two teammates who are not connected to one another. Figure 2 shows the number of structural holes to which each assigned position has access, which decreases as network constraint on the position increases. People assigned to position 5 are all subject to 77 points of constraint. Everyone assigned to position 1 is subject to 25 points of constraint. There is zero variation in treatment constraint on participants assigned to the same network position.

Behavioral Network During the Experiment

We recorded the frequency with which teammates messaged one other. The strength of the behavioral connection from teammate *i* to teammate *j* is the number of messages *i* sent to *j* during the experiment.[4] Since participants are often prohibited from contacting certain teammates in treatment networks, constraint in a participant's treatment network is strongly correlated with constraint in the participant's behavioral network (*r* = .964 , *N* = 465 participants).

When treatment permits variation, however, it can be pronounced. The more contacts a participant is allowed by treatment position (see Figure 2), the more his or her behavior *can* deviate from treatment. For example, participants assigned to one

---

[4]Measuring behavioral connection by message length (number of characters), yields constraint scores similar to the scores obtained with message frequency (*r* = .99). Message frequency and length both follow learning curves of many, long messages in early trials, progressing to fewer, shorter messages in final trials. There are four participants who made and received no messages. We treat the four as isolates, giving them maximum constraint scores, as they would be treated in a network analysis of managers. All four were assigned to a maximum constraint position in a treatment network, so having maximum constraint in their behavioral networks is no deviation from their treatment and keeps in the analysis their four separate teams.

of the broker positions are by treatment expected to operate under 68 points of constraint (Figure 2, position 4). But the behavioral network they experience differs. The experienced constraint varies from a minimum of 35 points to a maximum of 88 points (*SD* = 12) across participants. Constituents (assigned to position 6) also differ depending on who wins the broker rivalry (behavioral constraint varies from a minimum of 57 points to a maximum of 100 points). Similarly, there is high variation in behavioral constraint across participants assigned to the CLIQUE treatment network, ranging from 61 to 100 points (*SD* = 7.3). The high variation in experienced constraint indicates that participants assigned to the same position do not necessarily experience the same network.

Dependent Variable: Team Leadership and Performance

A central dependent variable here is the number of leader citations each participant receives over the course of the experiment. Leadership nominations are requested from each participant at three points in the experiment; after the 5th, 10th, and final trials (see Figure 1). The question: "Did your group have a leader? If so, who?" The response options were the color ID of the other people on the team (e.g., Purple, Black, Blue, Green, etc.), "Myself," "We worked as a team," or "Our team did not have a leader." The final survey also included a screen of background information on the participant (gender, age, education, and a textbox for any feedback the participant wanted to share about the experiment). If a team quit the experiment before the 15th trial, teammates were asked to complete the final survey before they left.

Given five teammates, asked three times, a participant could receive up to 15 citations as team leader. The final result is a skewed distribution, much as leadership is observed in organizations. A few people receive many leader citations. Many people receive few citations. Counts go from zero to 15 around a mean of .72 (*SD* = 2.28). Since teammates could only cite themselves or a direct contact as leader, an important control is the number of teammates eligible to name a participant as leader (five for participants assigned to positions 1, 4, and 5 in Figure 2; four for assignees to position 2, three for assignees to positions 3 or 6, and two for assignees to position 7). Since teammates could chose to name the whole team rather than an individual, an important control variable is the number of leadership citations available in a team. Counts of team citations available vary from zero to 15 (*SD* =

4.11) around a mean of 3.60. A person who receives three citations when 15 are available is less obvious a leader than a person who receives three citations when only five are available. For simplicity, we often control for available cites by predicting the percentage of available cites a participant receives. Percentage of cites is computed by dividing number of cites received by number of cites available and multiplying by 100 (setting to zero any participant in a team within which no one was cited as leader).

Beyond perceptions of leadership, the online experiment provides concrete measures of team performance in terms of speed, activity, and accuracy. We know the number of minutes that pass between the start of a trial and all teammates submitting their best guess of the shared symbol (4.54 minutes on average across teams and trials). We know the number of messages exchanged during that time period (65.93 messages on average across teams and trials). We know the proportion of guesses that were correct at the end of each trial (77.56 percent on average across teams and trials). All three performance measures display learning curves of weaker performance in early trials followed by increasingly good performance in later trials. Therefore, a fourth performance measure is how steep the learning curve is for a team. Some move quickly past weak initial performance. Others linger.

With a quick sketch of the experiment in hand, we turn to three exhibits illustrating the value of bringing phantom networks into the analysis.

EXHIBIT 1: GENDERED BROKERAGE

Evidence has accumulated on the benefits to people who broker connections in social networks — greater creativity, more positive work evaluations, higher compensation, and more rapid promotion than peers (Brass 2022; Burt 2021; Halevy et al. 2019; Kwon et al. 2020; van Burg et al. 2022). Contemporary interest in diversity, equality, and inclusion raised questions about unequal access to the advantages of brokerage. Unequal access by gender was an early concern. There are multiple ways to argue that returns to brokerage are gendered (Brands et al. 2022 for general discussion). Arguments are of two broad flavors: an audience effect and a production effect. The audience effect locates inequality in acceptance: Men are more readily recognized and accepted as brokers (Brands and Kilduff 2014; Burt

1998, 2010; Mehra et al. 1998). The production effect locates inequality in assertion: Men are more comfortable operating as brokers (Brands and Mehra 2019; Vedres and Vasarhelyi 2019). Burger and Buskens (2009:74) underline the relational hinge: “the existence of a broker is dependent on other actors who do not want to be brokers.” Related meta-analytic evidence is consistent with a gendered pattern in leadership emergence (Badura et al. 2018), but it rests largely on self-reported agentic traits (drawn from standard personality inventories [e.g., BSRI, PAQ]) and others’ ratings, so — as the authors caution — it cannot adjudicate production versus audience.

Much of available evidence is cross-sectional, nonexperimental, population-specific (this or that organization, as with much of management research), and sometimes based on no more than vignettes about hypothetical situations — which are open to reactive bias on socially charged topics such as gender and inequality – and rarely behavioral. Fang et al. (2021) look for behavioral corroboration. Their meta-analysis shows that men are more likely to hold brokerage positions, and their supplemental study links this to self-reported intention to build new ties for access or career gain. The effect is behavioral in name, but intentional in method — they ask what people meant to do; performed agency is unknown.

Behavioral networks offer a look at what people do inside the channels they use. We have participants randomly assigned to a treatment network and asked to work with randomly assigned teammates to solve coordination problems. Assignment fixes the formal channels, and time is zero-sum within a team; more attention to one connection means less attention to others. Interpersonal behavior during the experiment shows how participants use their prescribed channels. There is no discussion of gender in the experiment materials, and the gender of invisible teammates is unknown — muting an overt audience effect. A participant might be led by his or her gender bias to guess teammate gender, but it is no more than a guess. If production is a central driver of gendered brokerage, men should more often be the ones who step up to claim leadership, leaving teammates to move into, or drift into, other positions. Because gender is not disclosed and is not made part of the task, any gender difference in enacted brokerage is difficult to explain as an audience response. That moves the conversation from how gender shapes network outcomes to how it shapes network construction under constraint.

An Illustration

Figure 4 displays one team for illustration. The team was assigned to a CLIQUE treatment network, in which all participants occupy the same position (Figure 2, position 5). Yet one teammate (square) became the focus of attention through a high volume of messaging. Connections between the other teammates atrophied for want of attention. (Color coding in the experiment is augmented here by shapes to accommodate black and white printing.) Line width indicates the volume of messages between teammates during the experiment. In Figure 4, thick lines converge in the square participant (59% of all team messages are with the square), whereas thin lines connect most other pairs of teammates. The square participant, a man, is the broker in the team, and all teammates but one vote for him as the unofficial team leader. He connects with everyone, and teammates are less connected with each other than they are with him. He describes this as taking charge: “I eventually took charge and labeled everything so we could have a system. I would say leading and organizing was my strength.” His teammates — slightly older, educated, and evenly split between men and women — use more submissive language to describe their role in the experiment.

——— Figure 4 About Here ———

Measuring behavioral deviation, the parenthetical numbers in Figure 4 show how much each person’s position in the behavioral network deviates from their position in the assigned treatment network (participant constraint in behavioral network during experiment minus participant constraint in treatment network). Negative numbers indicate constraint lower in the behavioral network, and thus, a person who wiggled free of constraint in the treatment network. The person indicated by the square in Figure 4 lowered his constraint by five points (76.6 in the CLIQUE treatment network, 71.6 in the behavioral network). Constraint on him is less because his contacts are less-strongly connected with one another than they were prescribed to be in the treatment network. The 8.3 next to the triangle shows that his behavior resulted in increased constraint on another teammate (76.6 points in the CLIQUE treatment network, 84.9 in the behavioral network). Triangle’s strongest relationship is with the person marked by a square. Other teammate messages are also concentrated in the square.

Who Becomes the Emergent Broker?

Armed with a measure of behavioral deviation, we can watch a CLIQUE become a WHEEL. We can watch a broker emerge. We use that trace to ask who is more likely to make that transition — men or women? In our data, men are more likely to emerge as network brokers. We have 23 teams of five people assigned to CLIQUE treatment networks, for a total of 115 people, of whom 58 claimed they are men (in our exit interview or their Prolific profile). The other 57 participants are 44 who claim they are women, and 13 who leave the gender question blank or claim they do not identify as male or female.

On the behavioral deviation measure, men's behavioral networks are significantly less constraining than women's (-2.97 difference, t = -2.34, p = .02). The difference increases to -4.41 points when we compare men to all 57 participants who did not claim a male gender (illustrated by the -5.0 in Figure 4; t = 2.96, p = .004). In sum, men are more likely to emerge as network brokers.[5]

These mean differences are the average trace. The exhibit turns on the emergence event. When a man emerges as a broker leading the team, other men on the same team variably accept increases in constraint. A correlation between men and brokerage behavior supports gendered brokerage, but the core question is whether men tend to more often be the emergent brokers.

Figure 5 displays the distribution of behavioral deviation across participants assigned to CLIQUE treatment networks. The vertical axis is constraint in a participant's behavioral network minus constraint in the network position to which the participant was assigned for treatment (rounded down to the nearest integer absolute value, with extremes truncated to ±14). On average, the bars at negative

---

[5]Within the limits of our data, we explored other attributes associated with behavioral deviation, but gender is the primary correlate. For example, higher status participants should feel more at liberty to "guide" team discussion as they do in their personal lives (e.g., square in Figure 4). We have two status indicators, education (post-college, college, high-school and other), and employment (full time vs other). Behavioral deviation is independent of education (F(2, 100) = 0.21, p = .81), and negligibly associated with full-time employment (t = -1.92, p = .06), though there is a pattern of higher constraint for employment statuses other than full-time (4.90 average deviation for "part time," 2.38 for "student," and 1.37 for "unemployed and other"). We do not have data to explore the employment association, but the glimmer we see is consistent with a status story and invites further exploration. We also checked for the effect of more experience as a participant in Prolific tasks, but behavioral deviation is independent of the number of Prolific tasks a participant has had approved (t = -0.39, p = .70). Similarly, deviation is independent of participant race (t = 0.56 for white vs other, p = .58) and age (t = 1.78 across age, p = .08). Results in this footnote are based on robust test statistics.

values are greater for men, and greater in positive values for participants not claiming a male gender. That is the reason for the statistically significant negative correlation between male and behavioral deviation.

——— Figure 5 About Here ———

More specifically, the bottom of Figure 5 shows the gender mix of participants who emerge as brokers. Of the 23 teams assigned to CLIQUE treatment networks, 17 end up led by a single broker.[6] Those 17 are the behavioral deviation scores of -2 or lower. Of the 17 emergent brokers, all but two of whom are men. In other words, men are 88% of the emergent leaders with a monopoly on brokerage within the team. That concentration is the pattern. Even with so few observations, that is clear evidence supporting the production argument in gendered brokerage. If gender were in fact independent of network brokerage, the probability of seeing these results would be about two in a thousand ($\chi^2(1) = 9.96$, $p = .002$). Of course, a skeptic may still wonder whether text messages have inadvertently revealed gender, thereby reintroducing an audience effect. This seems unlikely. In seven teams, we asked at the end of the exit interview participants to guess the gender of their teammates. Six of 35 participants guessed the gender of every teammate, without being unanimous about anyone. The other 29 participants typically said that they did not have enough information to guess.

That said, some people are more likely to take the opportunity to emerge as brokers within their team. In teams of equals, men more often assert themselves, reshape the network, build bridges, and emerge as central brokers. Agency is situational in its trigger but gendered in its distribution. The asymmetry observed here is not explained by visibility, role, or attribution. It is the structure they build when personal attributes are invisible. The pattern echoes prior findings. Meta-analytic work shows men are more likely to emerge as leaders even without formal authority (Badura et al., 2018). Ridgeway (2001) explains how these outcomes become institutionalized: status expectations compound across interactions, shaping who acts and who defers. We do not lean on these literatures to explain our data, but

[6]The six teams in which a single broker did not emerge as leader typically had dual leadership in that brokerage was shared between a greater and a lesser broker (1 male pair, 1 female pair, and 4 mixed gender pairs with mixed levels of concentration in the brokers). With behavioral communication so often concentrated in a subset of teammates, one wonders how often CLIQUE treatments in past network experiments have been experienced as WHEEL structures.

they provide scaffolding for interpreting gendered brokerage as a pattern that is structurally emergent but sociologically familiar. We also note that our results do not deny a parallel audience effect — which could amplify or moderate the pattern. What behavioral deviation provides is a metric for introducing such audience effects in a controlled manner, isolating their influence from the documented production effect.

Gender Affects How the Team Works, Not How Well

Given evidence of a gender bias in who pushes to become the network broker, are there implications for team performance? Outside the experiment we see the job title, not the move that produced it. Within the experiment, we can watch the move and then ask about results. In practice, we expect consequences for individuals on a team since the person who takes the broker position is recognized by teammates for leadership, which likely extends to supervisors providing evaluations, compensation and promotions to people perceived to be leaders on teams in organizations (see reviews in the first paragraph to this section). The individual benefits we know, and the results here imply that men are more likely than women to receive brokerage benefits.

The question is whether the gender of a participant seizing the broker-leader role affects team performance — either improving performance through centralized leadership, or eroding performance through the struggle for leadership. The quotes in Figure 4 show no teammate irritation with the black square participant who emerged as team leader. However, not all texts were equally accepting of a subordinate role, and some participants simply quit.

——— Table 1 About Here ———

The results in Table 1 show that team performance is independent of who takes the broker role. As teams gain experience with the coordination task, they finish tasks faster (one task per trial), with less effort, and higher accuracy. The 23 teams assigned to CLIQUE treatments finished a total of 227 trials. From an average of 9.39 minutes on the first trial (intercept in first column of Table 1), teams did their third trial 3.41 minutes faster (-3.11 coefficient for ln[trial] in first row, first column in Table 1 multiplied by 1.098 change in ln[trial] from trial one to trial three). The first row in the table shows that, across trials, teams become faster (t = -7.36, p < .001), use fewer messages (z = -8.47, p < .001), and offer more accurate solutions (z = 2.98, p = .003).

The next three rows of the table ask whether broker gender and behavioral deviation from treatment shift performance. Results in the second row ask whether learning curves are higher or lower because of team communication concentrated in a single teammate. Tests in the third row are for learning curves being steeper or less steep because of team communication concentrated in a single teammate. Tests in the fourth row are for the gender of the teammate in whom communication is concentrated. The summary tests at the bottom of the table show that team learning curves are independent of all three behavioral deviation predictors.

We conclude that in a team with no formally assigned leader, someone stepping up to take the broker role as an informal team leader constitutes a change in how the team works but is independent of how well the team works (relative to other teams). This does not rule out offsetting effects that are unmeasured in the current design. To be sure, a next step would be to bring audience into the picture, to see whether visibility reshapes performance dynamics. Until then, this exhibit shows that gendered brokerage can emerge endogenously even when gender is not observed, an effect that, in this design, reflects production not audience.

EXHIBIT 2: RIVAL BROKERAGE

People can become bitter rivals in their efforts to appeal to the same constituency. Network theory grounds that rivalry, and competition more generally, in structural equivalence. Two people or groups are structurally equivalent when they have similar relations with the same other people. Structurally equivalent people are substitutes for one another in competing to be attractive partners in the same relationships. This is the foundation for the familiar competition between businesses in the same market (e.g., Burt and Carlton 1989; Burt and Talmud 1993), and people in, or aspiring to, the same role (e.g., Burt 1987; Gould 2003; Bothner et al. 2007; Burt 2010: Chap. 8; Piezunka et al. 2018; Kilduff et al. 2024). When the equivalence set reduces to two actors, competition can intensify into rivalry (Kilduff 2014; Kilduff et al. 2016; Kilduff et al. 2024).

When two brokers share a constituency, and thus are structurally equivalent, the contest over coordination sharpens because each has the same structural claim to lead, each a frame of reference for the other’s social standing. Rival brokerage refers to the competition that can develop between structurally equivalent brokers such that one works to dominate the other. As with gender, asking people about their competitors is fraught with emotional and status implications. Much of the available evidence is population-specific (e.g., Porac et al. 1995; Burt 2010:274-276) and efforts to make general inference are often forced back to vignettes about hypothetical situations (Martin et al. 2024). As with gender, behavioral deviation is useful to make visible where rivalry appears, how it unfolds over time, and when it fails to materialize.

An Illustration

Figure 6 displays the behavioral networks within two teams assigned to a CB treatment network (third treatment network from left in Figure 2). The participants indicated by a triangle and a pentagon are structurally equivalent brokers in that they both have connection with each other and the other three participants in the team. Conditions in the CB treatment network are ripe for broker rivalry: There is no defined difference in rank between the brokers (ambiguous status; Gould 2003) and there are only two, so their structural equivalence lends itself to rivalry.[7]

Line thickness in Figure 6 indicates relative message volume. In both teams, the thickest lines concentrate on one of the paired brokers. Message flow runs through one broker, not both. We will refer to the asymmetry in paired brokers as a distinction between the “lead” broker versus the “other” broker. We identify the lead broker as the one with lower constraint in the behavioral network. The lead broker behaves more like a broker during the experiment — ties to teammates thicken while ties among teammates thin. Both shifts reduce the lead broker’s constraint, yielding negative behavioral deviation.

——— Figure 6 About Here ———

In Figure 6A, the person assigned to the triangle position becomes lead broker. All teammate leadership citations go to the triangle. His constraint score in the behavioral network is 12 points lower than in the treatment network — visible in the

[7]There are other structural equivalent positions in the treatment networks, but not with the same potential for rivalry. The three subordinates in Figure 6 are structurally equivalent in their similar dependence on the two central brokers, but there is no direct contact between subordinates. The four peripheral members in a WHEEL network (position 7 in Figure 2) are structurally equivalent in their similar dependence on the broker in the central hub position, but they have no contact with one another. All five members in a CLIQUE network (position 5 in Figure 2) are structurally equivalent and aware of one another, but no two are distinctly structurally equivalent. The incentive for behavioral deviance in a CB treatment network is indicated by the high variance in behavioral constraint for both positions in the network (positions 4 and 6 in Figure 2).

thick behavioral lines that connect him with teammates and the thin lines between teammates. The pentagon position in Figure 6A is the “other” broker. Her constraint score in the behavioral network is 12.8 points higher than her score in the treatment network — visible in the relatively thick behavioral lines between her teammates. Note the dominant language used by triangle to describe his behavior during the experiment: “I really liked this. I was able to push and pull data from every person pretty well.” Illustrating the opposite side of rivalry, the pentagon “other” broker opines: “I felt like one team member got a bit bossy;” presumably referring to the person in the triangle position.

Figure 6B shows the same pattern in another team. All teammate citations for team leadership go to the triangle position. Thick lines are concentrated in the triangle. Constraint on the triangle in the behavioral network is 15.2 points lower than in the treatment network. The person assigned to the triangle position clearly felt his exercise of leadership, commenting: “My strength was analyzing the results and getting people to trust my conclusions.” On the other side of rivalry, the person in the pentagon position receives no citations as team leader, suffers an increase in behavioral network constraint of 19.1 points, and describes for himself a diminutive role in the team: “I focused on relaying messages quickly to double check people’s chats.”

When Equivalence Becomes Rivalry

The example teams in Figure 6 are not atypical. When paired brokers are connected, one of them usually ends up dominating the other. One broker’s freedom is the other’s constraint.

Direct contact turns out to be essential to rivalry. We have two treatment networks defining paired brokers: The CB network in Figure 6, and the DB network (second from left in Figure 2). The only difference between the two is connection between the paired brokers: yes in CB, no in DB. The paired brokers are structurally-equivalent competitors in both networks, but in the DB network, each can only infer a competitor from communication with subordinates. In the CB network, each broker’s competitor is obvious from direct contact. Rivalry would not be surprising in either network, but it should be more visible between the paired brokers in a CB network because both are more likely to be aware of their structural equivalence.

Ninety-four participants were randomly assigned to paired-broker positions: 46 to CB networks, 48 to DB networks. Figure 7 displays mean behavioral deviation by brokers in the two treatment networks. Note the sharp distinction between lead and other broker within CB networks. On average, the lead broker decreases constraint (average behavioral deviation score = -9.57), increasing constraint on the other broker (average behavioral deviation score = 7.02). The separation is sharper in CB than in DB, consistent with greater mutual awareness of equivalence, when the brokers are directly connected. Lead brokers are also more active in sending and receiving messages with teammates — an average of 39.7 messages per trial versus a lower 26.5 for other brokers.[8]

In all but two of the teams assigned to a CB treatment network, the lead broker's freedom comes at the cost of higher constraint on the other broker. Behavioral deviation by one is correlated at $r = -.81$ with behavioral deviation by the other. More specifically, each point of constraint avoided by the lead broker comes at a corresponding cost of .98 additional points of constraint on the other broker. Figure 6 shows what that looks like: One lead broker brags about pushing and pulling information among teammates, while the corresponding other broker complains that one team member got a bit bossy. The two exceptions are CB teams that expired early — one in the first trial, the other in the second — in an experiment composed of 15 trials. Why put up with a bullying broker when exit is an option?

——— Figure 7 and Table 2 About Here ———

One is reminded of Triplett's (1898:533) early work, often deemed the first empirical experiment in social psychology (Stube, 2005): "the bodily presence of another contestant participating simultaneously in the race serves to liberate latent energy not ordinarily available." Here the issue is not mere co-presence. It is direct connection. Connection is where rivalry sharpens to a point. As in Triplett's experiment, we do not ask participants about competition. We watch their behavior to see whether they make special effort when the situation includes an obvious competitor.

---

[8]Difference in mean messages per trial between lead and other brokers in CB networks: $t = 3.08$, $p = .004$ ($N = 46$ paired-broker participants assigned to CB networks). The association between paired brokers' behavioral deviation in CB teams implies a near one-for-one tradeoff: coefficient -0.98 ($t = -9.00$, $p < .001$).

In contrast, there is a relatively small difference in Figure 7 for behavioral deviation between the lead and other brokers in DB networks (mean behavioral deviation scores of 1.00 and 2.07). The paired brokers are also equally involved in messaging with non-broker teammates (respective averages per trial of 36.5 and 34.8; t = 0.39, p = .70). More important, their behavioral deviation is aligned rather than opposed. In contrast to the negative correlation in CB networks — one broker becoming less constrained while the other becomes more constrained — the correlation is positive in DB networks (r = .92). For every additional point of constraint on the lead broker, the other broker acquires an additional 1.53 points of constraint (t = 18.69 , N = 48, p < .001). As an indirect indicator of reduced tension in DB networks, these teams are more likely to claim that leadership was exercised by the team as a whole rather than by specific individuals (4.44 loglinear test statistic, p < .001). Sixty-one percent of participants in CB networks made that claim. Eighty percent of participants in DB networks made the claim, which is even higher than the 73% of participants in CLIQUE networks, wherein all were prescribed to be equal teammates.

The regression results in Table 2 put summary numbers to the contrast between paired brokers who are connected (CB) versus not (DB). Because “lead broker” is defined within each team as the broker with lower constraint in the behavioral network, Table 2 necessarily shows a small but statistically significant decrease in constraint in the behavioral network for lead brokers, and a corresponding increase for the other broker. One broker more dominates the other in CB networks (t = -5.83, p < .001). In other words, rivalry appears between structurally equivalent brokers when they are able to communicate with one another. Without direct contact the behavioral trace of rivalry disappears. Mere inference of a competitor is not enough. Rivalry only emerges when structural equivalence is made actionable through direct contact between the brokers.

Recall that audience cues are muted in this experiment, so the dominance that appears in CB teams is a production effect, not performance signaling. This does not mean that status is generally irrelevant. Cao and Smith (2021) show how visible status can broaden networks by increasing expected receptivity. Carnabuci et al. (2018) show that local hierarchy can emerge in informal groups absent formal authority. Our rival-brokers findings point to one such route: Through repeated interaction, one broker builds a stronger constituency and turns structural equivalence into local dominance.

Quick aside on the gendered-brokerage exhibit: Rivalry dominates gender as a correlate of broker emergence in these teams. Figure 7 shows that men are slightly more likely to be the lead broker in CB networks. Among participants assigned to broker positions in CB networks, 49% of the lead brokers are men, compared with 39% of the other brokers. That result echoes the earlier CLIQUE result: When teammates have equal access and no one is assigned to lead, men are more likely to emerge as broker. With respect to rivalry, however, gender adds nothing to the prediction in Table 2. Gender is independent of behavioral deviation here. This is true whether we add a dummy variable for participants identifying as men (t =0.40, N = 94, p = .69), whether we limit the dummy variable to participants as men versus women (excluding “other,” t = -.46, N = 85, p = .64), or whether we simply test whether men tend to be lead brokers in CB networks, where the gender difference is most apparent ($\chi^2(1)$ = 0.35, N = 46, p = .55). The counts are small, but the pattern is consistent: Rivalry affects women as it affects men. In CLIQUE networks, where teammates have equal access to one another and no one is assigned to lead, men are more likely to step up. But in CB networks, where structure creates rivalry, men and women are equally likely to take the lead. These results could change, of course, if paired brokers were aware of one another’s gender. Here, one’s rival is unseen and unheard.

Process Trace Across Trials

We have used behavioral deviation as an agency diagnostic: Who becomes lead broker, and who the other? Now we illustrate its value as a process trace: Recorded communication lets us see when rivalry begins, how it consolidates, and when it never takes hold. Figure 8 contains plots of message activity by trial for teams assigned to CB treatment networks. The experiment runs for 15 trials, each trial a coordination task to be solved. All three graphs show teammates learning to work together. Message volume is high in early trials and declines as teams learn to solve tasks with fewer messages. Figure 8A shows message activity by participants who emerge as lead brokers. Lead brokers send messages to the other broker (solid line) and to any of the three subordinates (dashed line). The graph shows a mean of 8.0 messages to the other broker during the first trial, and a mean of 10.3 messages to each subordinate. (To make the lines comparable, the total count of messages to all three subordinates is divided by three so that the solid and dashed lines reflect messages per teammate per trial). These message counts drop to an average of 2.5 messages per trial by the end of the experiment.

——— Figure 8 About Here ———

In parentheses, we average message counts within three stages of team development: novice, intermediate, and mature. For example, participants who emerge as lead brokers send to each subordinate an average of 8.4 messages during the novice stage, 5.1 during the intermediate stage, and 2.5 during the mature stage. The initial four trials define the novice phase, when teammates perform least well and rapidly learn how to work together. The team enters an intermediate phase from the fifth through the tenth trials. Learning continues, but more slowly than during the novice phase. On average, teams reach the mature phase after the tenth trial, indicated by a lack of improvement. The curves in Figure 8 are relatively flat across the last five trials of the experiment.

Lead brokers tend to emerge in the initial trials. Team activity begins with the person who will end up lead broker sending more messages to the subordinates than to the other broker (8.4 versus 6.8 in Figure 8A—a statistically significant difference indicated by the asterisks next to 8.4).[9] In contrast, the person who will end up the "other" broker sends the same number of messages to all teammates (5.5 in Figure 8B).

In the intermediate phase of team development, the lead broker sends more messages to each subordinate than to the other broker, but the margin narrows relative to the early trials (5.1 to each subordinate versus 4.5 to the other broker). In contrast, the person who will end up the "other" broker now sends more messages to

---

[9]To provide a sense of larger message frequencies, statistical tests are from a regression model predicting average number of messages a person sent to a kind of contact during a trial (during a phase of team development). For example, there are 162 trials during the intermediate phase of team development in which lead brokers sent messages to the other broker and subordinates. Here is the regression model: Y = 12.92 - 4.33(ln trial) + .58 X, where observations are a sender-trial, X is a dummy variable equal to 1 for messages to subordinates and 0 for messages to the other broker, and Y has two values for each sender; average number of messages a lead broker sent to subordinates and average number of messages the lead broker sent to the "other" broker. Estimated across 186 observations, routine test statistics are *t* = -4.43 for the learning curve predicted by ln(trial), and *t* = 1.26 for the more frequent messages to subordinates. The 1.26 fails to reject the null (*p* = .21), so there are no asterisks in Figure 8A on the 5.1 average number of messages sent to each subordinate during the intermediate stage of team development.

the lead broker than to the subordinates (3.5 messages versus 2.7, a statistically significant difference), and continues to be less active overall. By the intermediate phase, the roles are established. While the lead broker typically initiates dominance through early message volume, we cannot rule out a small Matthew Effect in the first trial — early attention from teammates that the broker amplifies through active response.

When the team has settled into a routine during the final trials, the lead broker sends the same small number of messages to subordinates and to the other broker (2.5 and 2.4, respectively), and the other broker sends a similarly small number of messages to both (1.9 and 1.9, respectively).

Turning to Figure 8C, subordinates have fewer contacts with whom they can exchange messages (two versus four for the brokers) so they end up sending more messages per teammate (the lines sit higher than in Figures 8A and 8B). The key point is that subordinates, in the initial trials, are responding to the higher message volume from the person who will emerge as lead broker: 9.1 messages per trial to the lead broker during the novice phase, versus 7.6 to the other broker, and 5.3 per trial during the intermediate phase versus 4.2 to the other broker. Both differences are statistically significant beyond a .001 level of confidence. When the team has settled into a routine in the final trials, subordinates send equal numbers of messages to both brokers.

In sum, the process trace shows that rival brokers do not attack one another so much as compete for constituency among shared subordinates. Eighty-two percent of all lead-broker messages are to subordinates; eighteen percent are to the other broker. Focusing on subordinates in early trials strengthens connections between contacts in the other broker's behavioral network, increasing constraint on that broker. The broker who invests more in subordinate ties emerges as lead broker during the novice phase—a role acknowledged by the other broker, who sends more messages to the lead broker than to subordinates during the intermediate phase. That is the leverage: Watching deviation unfold reveals how structural equivalence plays out in action. Rivalry manifests as constituency-building. Dominance is the outcome.

Rivalry Affects How the Team Works, Not How Well

Because of the tensions associated with rivalry, we expected rivalry between brokers to erode team performance. It did not. The results in Table 3 show that team performance is independent of broker rivalry. The teams assigned to the two paired-broker treatment networks completed 458 trials. Teams begin slowly — an average of 9.81 minutes on the first trial (intercept in the first column of Table 3) — but improved with experience. By the third trial, teams are completing tasks 3.44 minutes faster. Across trials (row 1), teams became faster (*p* < .001), used fewer messages (*p* < .001), and produced more accurate solutions (*p* < .001) — as illustrated by the learning curves in Figure 8. That pattern holds across both treatment networks.

——— Table 3 About Here ———

Results in the next two rows of Table 3 show that team performance is stable across rivalry conditions. Teams assigned to rivalrous CB networks perform no better or worse than those in DB networks. Learning curves are neither higher nor lower (second row), nor steeper or flatter (third row), for teams assigned to CB networks.

The same holds for rivalry intensity. Team performance does not vary by the magnitude of behavioral deviation between the lead and the other broker. Teams with sharp behavioral asymmetries do not learn faster or slower, and their learning curves are not steeper nor flatter than for teams with less behavioral deviation.

Summary tests at the bottom of Table 3 confirm the pattern: Team learning curves are independent of all four network predictors above.

We conclude that rivalry affects how a team works — what it feels like to be inside — but not how well the team performs. Prior research links brokerage to individual rewards, and it is reasonable to assume that rivalry affects recognition of individual performance when one broker emerges as the leader. By extension, it is reasonable to expect broker rivalry to affect team performance. But here, rivalry leaves no mark on speed, effort, or accuracy. As with gender, we observe rivalry as structure, not attribution. Because audience cues are muted by design, our evidence is sharper on rivalry over coordination within a shared constituency than on rivalry as status display. Brokers emerge through behavior, not role or recognition. That may also explain the null performance effect: stripped of status cues and audience response, agency is visible, but its social perception is irrelevant.

EXHIBIT 3: STRUCTURAL REPRODUCTION

For our third and final exhibit, we turn to interpersonal behavior outside the experiment as a phantom inside the experiment. Using the generic General Social Survey instrument, we have data on each participant's personal discussion network before entering the experiment. The histogram in Figure 1 shows that our participants span a wide diversity of networks. Some come into the experiment with extremely open networks (frequent discussion with several people who have little contact with one another; near-zero network constraint score). Some enter the experiment with extremely closed, cliquish networks (frequent discussion with few people who have frequent contact with one another; near-100 network constraint score). The distribution is roughly bell-shaped around a 55.64 mean constraint score (*SD* = 16.61).

Large differences in achievement are associated with such network differences. Holding constant differences in job rank, kind of work, and related considerations, people with more open, less constrained discussion networks are more likely to come up with good ideas, receive positive work evaluations, enjoy higher compensation, and receive faster promotions (Brass 2022; Burt 2021; Halevy et al. 2019; Kwon et al. 2020; van Burg et al. 2022). These people are often discussed as a type of person — a connector, a cosmopolitan, a network broker — who is comfortable making connections across otherwise disparate groups. At the other extreme, the low-achievement people in closed networks are discussed as an opposite type of person. Implicit in such discussion of types is the assumption that the personal network around a person indicates how a person has learned to behave interpersonally, which is how we expect them to behave in future. This is an assumption of structural reproduction. If a person is familiar with a kind of structure, we expect that person to build the same structure again and again. Some people build open networks again and again. Some people huddle in closed networks again and again. There is research to cite in support of expecting structural reproduction. For example, Martin et al. (2024) show network closure persisting from high school into college (contingent on networks in the college selected). Kleinbaum and Stuart (2014) show closure persisting into later career (contingent on selection into a job at headquarters). More directly linked with our experimental setting, Janicik and Larrick (2005) show that people in open networks pre-experiment more quickly identify structural holes in treatment networks. All of which led us to expect in our experiment that participants in closed personal networks would be more passive recipients of the treatment network to which they were assigned. Participants accustomed to open personal networks we expected to wiggle free of constraint, creating a social environment within the experiment more like the environment in which they live outside the experiment.[10]

More Closed Outside, More Open Inside

People are assigned at random to positions in treatment networks, so differences in personal network are independent of treatment. There is no correlation between constraint in a participant's personal network and the treatment constraint to which they are assigned (r = 0.07, t = 1.62, p = .11).

But look at Figure 9. Personal networks are closely associated with behavioral deviation from treatment during the experiment. The graph in Figure 9 distinguishes participants across the horizontal by personal network outside the experiment, open networks to the left, closed to the right. The vertical axis measures behavioral deviation. High positive scores indicate a participant whose interpersonal behavior during the experiment resulted in higher constraint than was prescribed by the treatment assigned. These are the people who ended up the "other" broker under rivalry (pentagons in Figure 6). Strong negative scores indicate a participant whose behavior resulted in decreased constraint. These are the people who became the dominant broker under rivalry (triangles in Figure 6), or converted a treatment CLIQUE into a WHEEL in which they were the hub (square in Figure 4).

——— Figure 9 About Here ———

The negative association in Figure 9 rejects structural reproduction. It is not the people with open networks who wiggle free of treatment constraint. It is the people in closed networks. The more closed the network a participant brings into the experiment, the more likely the participant wiggles free of treatment constraint. Across all participants, there is a statistically significant negative correlation between behavioral deviation and personal-network closure (t = -2.04, p < .05). The

---

[10]There is evidence that investment bankers and analysts oscillate between the extremes of open and closed networks (Burt and Merluzzi 2016 on oscillation), but that result awaits replication, and suffers from the suspicion that such financial elites live in their own special world.

correlation is especially strong (r = -.85, Figure 9) among participants assigned to a treatment network prone to behavioral deviation: CLIQUE or CB networks (t = -5.61, p < .001). The negative association describes both men and women,[11] regardless of rival brokerage.[12] There is a hint of curvature in the graph, but close analysis supports a simple linear association.[13]

In short, behavior during the experiment is not independent of one's personal network outside the experiment, but neither does it reproduce the outside network. We observe a reversal. The behavioral deviation described in the last two exhibits does not come from experienced brokers reasserting familiar roles. Participants entering from closed personal networks — that is, cliques in which mutually-familiar discussion partners speak frequently with one another — are the most active in messaging teammates. Their activity crowds out other ties. Teammates have less time to message one another. The active participant becomes a broker — a person strongly tied to teammates who are weakly tied to each other.

Participants entering from open networks, more experienced with the diversity there, tend to be the ones displaced. Figure 9 shows pre-experiment brokers (low constraint, left side) driven into closed positions during the experiment (top-left side of the graph). Some experienced brokers do reassert brokerage, but most end up more constrained than treatment prescribed. Perhaps people more experienced with

---

[11]The aggregate correlation in Figure 9 is r = -.80 for men and r = -.63 for women. In a regression model across individuals assigned to CB or CLIQUE networks, there is no level difference in behavioral deviation for men versus women (t = 0.29, p = .77), and no slope difference for personal network mattering more for women in predicting behavioral deviation (t = 0.08, p = .93).

[12]We ran the regression model in Table 2 using as a dependent variable network constraint in a participant's pre-experiment network, expecting that participants more experienced with network brokerage might be more likely to identify the rival broker and so emerge as lead broker. There is no difference in pre-experiment network between the two treatment networks (because of random assignment, t = -1.46, p = .15), and no difference in pre-experiment network between participants who end up as lead versus other broker (F(2, 90) = 0.16, p = .85).

[13]As a reminder, t-tests here are based on Stata's "robust" standard errors. The first few data points represented by the left two averages in Figure 9 (hollow dots) show an upward sloping line followed by a downward sloping line across higher levels of constraint. To check for a nonlinear association, we used an analysis of covariance model in which behavioral deviation (vertical axis in Figure 9) is predicted by personal-network constraint (horizontal axis) plus a level adjustment for constraint lower than 30 points, and a slope adjustment for the x-y association slope being positive for constraint scores lower than 30 points. The t-test for the aggregate association with constraint is t = -3.06 (p = .002). The tests for the level and slope adjustments are t = -0.58 and t = 0.27 respectively, indicating negligible adjustments. The behavioral deviation of participants with extremely open networks is a thing to watch for in replication work, but for the purposes here, we put them aside to focus on the dominant positive association in Figure 9.

the diversity of open networks are more likely to wait to get a sense of the group before they launch into messaging.[14]

For some participants, particularly those entering from closed networks, the experiment may function as a context for temporary freedom — a space to act differently than they do in everyday life. Agency in that sense is not necessarily expressed through familiar network positions, but may be exercised opportunistically: behavior they seize the chance to try out. What we see is evidence of responsive initiative, not simply the reproduction of prior network structure — agency that emerges when the situation allows it. This supports a view of agency as partly situational, rather than fully determined by prior dispositions: not reducible to prior network socialization, but also shaped by local conditions, available roles, and perceived interaction opportunities — at least when audience recognition is muted. These results resonate with Burt's (2012) multirole analysis of agency. In his analysis, only a third of network variance carried across roles, while the larger share was situational and role-specific. Our evidence is a behavioral parallel, highlighting that brokerage can emerge from situational opportunity, revealed when structure gives room. If correct, structural stability may in part reflect recurrent situations that limit initiative, rather than internalized behavior carried from one setting to another. This contrasts with Burt et al.'s (2022) finding that closure carries behavioral consequences beyond the network in one-shot exchange; in our setting of repeated interaction within an assigned group, actors have more room for situational agency in response to local openings.

A check on our interpretation of agency here is to compare messages sent and received. If behavioral deviation varies more with teammate messages received, then teammate agency warrants attention. If behavioral deviation varies more with a participant's own messaging, then we feel safe saying deviation is driven by the closed-network people extensively messaging. For participants in the two treatments prone to behavioral deviation — CLIQUE and CB networks — we compare messages sent and received. Behavioral deviation has a strong negative association

[14]Our conclusion in the previous footnote is reinforced if we look at the people most responsible for behavioral deviation. Of the 23 lead brokers in CB networks (Figure 6), one is in a hollow dot to the lower-left in Figure 9. All the others are to the upper-left in the graph. Of the 23 people assigned to CLIQUE networks who emerged lead brokers (deviation scores of negative two or more in Figure 5), two are in a hollow dot to the lower-left in Figure 9. All the others are to the upper-left in the graph.

with the number of messages sent (t = -8.13, p < .001 holding constant treatment network and last active task) and with words sent (t = -4.16, p < .001, with the same controls). The more messages/words sent, the more that constraint in the behavioral network is below constraint in the treatment network. In contrast, behavioral deviation has a weak negative association with messages and words received (F(2, 169) = 3.08, p = .05). We are accordingly more confident that the behavioral deviation in Figure 9 comes from closed-network participants, who are more active in messaging teammates and therefore emerge as brokers.

Of course, pre-experiment personal networks matter to the extent that treatment structure allows for deviation. The correlation disappears in WHEEL and DB treatment networks, where participants are offered less choice. That defines a boundary condition: Personal networks shape behavior under porous treatments, but are inert in tightly-scripted treatments. Another condition here is that participants from closed networks outside an experiment might be especially sensitive to an audience watching their behavior. Here, interpersonal behavior is dyadic. Regardless, it is certainly the case that closed-network participants do not lack agency and initiative.

——— Table 4 About Here ———

Outside Network Affects How the Team Works, Not How Well

Despite the interpersonal struggles implied by the emergence of some participants as brokers, personal networks outside the experiment have no association with team performance. Teams come down their learning curve regardless of participants drawn from closed or open networks outside the experiment. Across trials, results in the first row of Table 4 show that as teammates work together, they become faster (p < .001), use fewer messages (p < .001), and offer more accurate solutions (p < .001). Results in the next four rows of Table 4 test whether the learning curves are affected by the most open or most closed personal network on the team. Summary tests at the bottom of the table show that team learning curves are independent of all four network predictors.

Behavioral deviation is independent of team performance in all three exhibits. Linking back to Coleman's (1990) boat metaphor, one could read this as evidence that variation in micro-processes matters little for macro outcomes. We draw a different conclusion. The value of observing behavioral deviation is not that it necessarily reveals micro-processes with large performance consequences. The value is that it makes those processes observable and therefor testable. In our data, substantial variation in agency, leadership emergence, rivalry, and behavioral adaptation is associated with little variation in team performance. But that result is itself an empirical finding, not a general implication of the framework. The broader contribution is a way to make the middle of the boat visible, providing a way to examine which micro-processes matter for macro outcomes, and under what conditions.

Before shifting to future research, a quick caution on generalization is warranted. One would think that gendered brokerage and teammate rivalry affect team performance. We find abundant evidence of effects on how a team operates, but no implications for team speed, effort, or accuracy (Tables 1, 3, and 4). How well a team works is predicted by network theory, and that theory can be wrong if it reasons from treatment rather than the network experienced. For example, the team assigned to a CLIQUE treatment in Figure 4 in fact experienced a WHEEL network because one teammate emerged as a broker coordinating contact between the other teammates. As predicted by network theory, the teammate who emerged as the broker (square in Figure 4) is the person who received all teammate citations as informal team lead. In general, the network of interpersonal behavior during an experiment will offer better support for network theory than treatment networks — because the behavioral network is the actual network experienced. That opens productive questions about testing network theory, but our concern here is only proof of concept for the value of capturing phantom networks to study behavioral deviation as a window on agency.[15]

[15]This is a subject pursued elsewhere, but here is a rough sense of gain from recording behavior during the experiment: Across the participants discussed here, an equation using constraint in the treatment network to predict leader citations received generates R^2 = .200. The same prediction using the behavioral network that develops during the experiment generates R^2 = .226, a 13% increase in predicted variance (Reagans and Burt 2026: Table 3). If prediction is limited to participants assigned to CLIQUE and CB networks — the two treatment networks in which behavior most deviates from treatment (Figure 2 above, Figure 10 below) — the respective values are R^2 = .119 and R^2 = .198, a 66% increase in predicted variance.

PARTING THOUGHTS

Our goal for this discussion has been to offer proof of concept for the idea that the behavioral networks often ignored in network experiments can be a valuable window on agency. We consider two such networks: the personal discussion network a participant brings into an experiment, and the message network that emerges during an experiment. Respectively behavior external and internal to the experiment, the two networks are phantoms in standard research practice, forms of noise and bias to be managed. We study that noise and bias as signal. We measure the constraint a participant experiences in each network, and the extent to which interpersonal behavior deviates from treatment — taking that behavioral deviation as a window on agency.

Having shown it is not difficult to obtain the proposed data, we offered three exhibits of agency revealed. The exhibits — on gendered brokerage, rival brokerage, and structural reproduction — do not propose new ideas. The topics are familiar in prior research, which is why we chose them. Studying behavioral deviation can help us more clearly understand the familiar by making agency visible.

In the first exhibit, gendered brokerage appears even when the gender of teammates is invisible. In equal-access networks (CLIQUE treatment), one person often takes over coordination, becoming the hub through whom messages flow (Figure 4). That person is often a man; not because men are assigned brokerage, but because they assert it during the experiment. The pattern is consistent with a production effect operating even when overt audience cues are muted. At the same time, our results do not deny a parallel audience effect, which could amplify or moderate the pattern when gender becomes visible. What behavioral deviation provides is a way to distinguish these influences by making the production side visible.

In the second exhibit, we watch rivalry develop. Rivalry does not develop between alternative leaders for the same constituency if each is kept in the dark about the other (DB treatment). However, if the leaders can communicate (CB treatment), rivalry develops, not in the form of each attacking the other so much as from one building ties with subordinates, pushing the rival into a secondary position, and emerging as the lead broker (Figure 6). Qualifying our first exhibit, rivalry dominates gender. The gender difference in who emerges as leader in teams of equals disappears when structure defines rivals. Agency is anchored in the person — male or female — who goes after the rival's constituency.

Finally, in our third exhibit, we watch structural reproduction occur in reverse. Network theory predicts that people carry forward familiar roles, re-creating positions they know. By that prediction, participants who enter the experiment with open networks should reassert brokerage when structure allows. Instead, we see the opposite. It is participants from closed networks who are more likely to deviate from treatment so as to make themselves lead broker (Figure 9). These "emergent brokers" send high volumes of messages to all teammates, crowding out communication between the teammates. People who enter with open networks in their outside life tend to end up more constrained than prescribed by treatment. This reversal runs against intuitive prediction and raises questions about often-cited sources of stability such as personality (Mehra et al. 2001) or broker experience (Burt 2005). The reversal is consistent, however, with images of social flexibility such as multirole evidence showing that only a small fraction of network behavior carries across settings (Burt 2012), or evidence of success contingent on oscillating between open and closed networks (Burt and Merluzzi 2016). Brokerage is asserted here less as habit than as an opportunity seized. Agency is situational, raising the possibility that structural auto-correlation reflects not durable habit but similar situations that leave little room for agency. The same situational logic is visible in our first and second exhibits: Women do not habitually avoid brokerage as might be inferred from our gendered brokerage exhibit. In the rival brokerage exhibit, they act no differently than men in dominating their rival.

Next Steps: Replication

We have reported on a secondary analysis of a series of experiments using adults in the United States as participants. The results, we believe, are sufficient proof of concept to warrant further work toward converting the behavioral networks into a concrete element in research design. However, the first order of business is replication. Do colleagues elsewhere — as a result of adding a GSS name generator to participant pre-work, or recording participant contact with teammates during an experiment — obtain similar results for participants from other study populations in their own variations on the iconic Bavelas-Smith-Leavitt team experiment? That work can move forward from the preceding discussion.

Next Steps: Structures

Our treatment networks include completely connected teams (CLIQUE), teams in which a leader coordinates disconnected subordinates (WHEEL), and two hierarchical teams (CB and DB networks). These networks cover a broad spectrum of structures, but do not include two networks widely consequential in organizations — project networks composed of two or more cliques (common in studies of alliance networks, e.g., Ahuja 2000; Koka and Prescott 2008), and partner networks composed of a broker who gains access to a target audience through a sponsor (Burt 1998; Burt 2010:Chap. 7). Both are logical extensions of the design considered here.

What about history? Burt and Opper (2024, 2026) document the competitive advantage of network brokers among Chinese entrepreneurs: Brokers enjoy higher profits from their ventures. But the advantage depends on personal events that years ago established relationships. Networks in this paper have no history. What if teammates were friends? We know that Joe is bossy, but we suffer his pushy behavior because of his other good qualities. Would teammates patiently suffer the emergent leadership of Joe as in Figure 4 and Figure 6, or would they actively seek one another out, thus reinforcing the CLIQUE structure of the network?

A more ambitious extension would allow participants to modify the treatment network itself by adding or deleting ties. Experimental designs that permit tie addition, deletion, or both could provide a valuable window on agency by making network construction visible, moving the question from how participants use structure to how they create structure. Similar opportunities arise in experiments where some contacts are algorithmic rather than human. Behavioral deviation could reveal whether participants respond similarly to AI partners or whether responses vary systematically with the personal networks they bring into the experiment.

Work exploring any of these questions can move forward from the preceding discussion. Implicit in that discussion, however, is a strong recommendation about the kinds of structures most likely to make agency visible. We want to make that recommendation explicit before we close.

Research intended to study agency is advised to focus on what can be termed “weak structures.” In the person-situation literature, “weak situations” are contexts in which behavior is less constrained and individual tendencies are more likely to emerge (Mischel 1968; Meyer et al. 2018). The analogy in networks, would be “weak structures” that encourage the exercise of agency.

The point is illustrated in Figure 10. The seven positions in our four treatment networks are arrayed across the horizontal. Behavioral deviation runs up the vertical. High positive scores indicate people whose behavior during the experiment resulted in networks more constraining than what was prescribed by treatment. Strong negative scores indicate people whose behavior allowed them to wiggle free from constraint prescribed by treatment. The figure is a box-and-whisker display of behavioral deviation (see note below figure title).

——— Figure 10 About Here ———

The three boxes to the right describe teams assigned to CLIQUE or CB treatment networks. These treatments can be termed “weak structures” in the sense that they allow considerable behavioral deviation. For precisely that reason these are unattractive in the traditional research-design sense of controlled treatment, but they are a productive source of behavioral deviation data on agency — which is why we use them in our exhibits.

The four boxes to the left in Figure 10 describe teams assigned to DB or WHEEL treatment networks. These treatments can be termed “strong structures” in that they show little behavioral deviation. Each position in these treatment networks is connected to disconnected contacts. For example, the hub of the WHEEL treatment (position 1) is connected to four teammates, none of whom is connected to the other contacts. Each of the paired brokers in the DB treatment (position 2) is connected to three teammates, none of whom is connected to the other contacts. The only discretion allowed in these treatments is to send extra messages to teammates that you find more appealing for whatever reason. When you do that, you increase the level of behavioral constraint on you because your relations become more concentrated in the chosen teammates. Thus, behavioral deviation for these networks is primarily positive, indicating higher constraint in the behavioral network than is prescribed by treatment. At the extreme are the poor souls assigned to a peripheral position in the WHEEL treatment (position 7, far right in Figure 10). Their treatment network gives them no discretion. They can send more or fewer messages to the hub, but the hub is their only option. Accordingly, they are maximally constrained behaviorally and in treatment, so they show no behavioral deviation at all.[16] The WHEEL and DB networks can provide a control baseline against which deviation in other networks can be compared, but one could just as well — and at lower cost — use zero deviation as a baseline.

Next Steps: People

All of our three exhibits display people emerging as network brokers, providing informal leadership to teammates. These people are men in the gendered brokerage exhibit, men and women in the rival brokerage exhibit, and men and women with closed personal networks in the structural reproduction exhibit. What can we say about the background or personalities of these people?

Our limited data show that gender is the primary predictor of who emerges as broker in CLIQUE networks (footnote 6), but gender is irrelevant to who emerges as lead broker under rivalry (Table 2), and irrelevant to people from closed networks more likely to emerge as brokers during the experiment (footnote 12).

People who have open networks are often reported as high in self-monitoring as a personality characteristic (e.g., Fang et al. 2015; Mehra et al. 2001), but those correlations are with survey network data similar to our General Social Survey measure of networks pre-experiment. Is it therefore the case that the correlation with self-monitoring is also reversed in the experiment — with low self-monitors more likely to emerge as the leader brokering connections among teammates? The negative correlation in Figure 9 begs for resolution with prior research on kinds of people associated with brokerage.

Next Steps: Scale

Shirado and Christakis (2017) show that coordination tasks can be performed more effectively by teams of 20 people if three people are replaced by robots generating a low level of randomness from a position central in the team network. They are able to employ teams of 20 because their coordination task does not require as much

---

[16]Participants assigned to this position voice frustration with their feet. They quit the experiment. People assigned to the CLIQUE or two paired-broker treatments drop out of the experiment at similar rates: 83% are still present after the novice trials, 47% are still present after the intermediate trials, and 29% are present through all 15 trials (trial categories are defined by stages in the Figure 8 team learning curves). People assigned to the WHEEL treatment are more likely to exit: 74% are still present after novice trials, 30% after intermediate, and only 9% survive through all 15 trials ($\chi^2(3) = 19.87$, $p < .001$) comparing four dropout categories for WHEEL versus other treatments.

communication as we do between teammates. What would happen to our results if our teams were 20 people instead of five? Even with teams of five, we have a large number of frustrated people quitting the experiment.[17] We expect that the odds of quitting increase with higher coordination difficulties in a larger team. In a CLIQUE network, for example, number of coordination ties increase roughly with the square of teammates (N[N-1]/2, where N is teammates). Thus, we expect a small-scale “tragedy of the commons” in which people quit or free ride in larger teams (Olson 1965; Hardin 1968). In an undifferentiated structure (CLIQUE), the size question is straightforward; run teams of five, seven, nine, etc. looking for the size after which teams cannot function. For what size teams do our results hold? Do larger teams need to be subdivided into subgroups to see agency? For example, are people with closed personal networks as likely to take over larger teams as they are likely to take over coordination in five person teams? In a differentiated structure, such as the CB network generating rivalry, an increasing number of subordinates might intensify rivalry, while an increasing number of alternative leaders could dampen it, increasing the odds of quitting.

Next Steps: Audience

A central feature of our experiment is that audience effects are muted. We did that to highlight production, to see what people are inclined to do before behavior is filtered by visibility, role expectations, or after-the-fact attributions. In much prior work, agency is inferred from respondent recall, self-assessment, or audience perception of what others did or appeared to do. Here we observe agency directly. Teammates are neither heard nor seen. Status is unmarked. Behavior unfolds without reputational consequence. The result is a record of agency unobscured by identity cues and the consequences that characterize organizational life. But to what extent does the emergence of closed-network people as team leaders in our third exhibit depend on the absence of colleague judgement?

Management behavior typically occurs with an audience watching, real or imagined. Emergent brokers can look more challenging than collaborative; less like people working with the team than people barging into it, imposing leadership as a solution to frustrating teammate behavior. Heavy messaging among co-workers can

---

[17]See preceding footnote.

increase visibility, trigger pushback, or draw alters closer together in reaction. Visibility, reputation, and attribution can amplify, mute, or redirect the behavior we observe. A challenge for our proposal is to understand how agency is amplified or erased by context. Do people respond differently when gender is visible, when they know the history of the contact, when merit is marked, or when origin is known? The same initiative may be read as leadership, competence, or presumption depending on who is seen to be acting and by whom. The question is not only how context constrains behavior, but how action becomes visible within it (e.g., Gulati and Srivastava 2014, esp. Table 1; Tasselli and Kilduff 2021). As Kwon et al. (2020:1110) conclude their review of brokerage research: "… there is need to disentangle opportunity, awareness, and intent." The audience we have removed is key to answering these questions. When and how do audience effects shape the fate of agency? In short, an important next step is to bring audience back in, one element at a time. Then we see what becomes of the agency we have observed. Bringing audience back is the most complex of our listed next steps.

We leave excited about the research potential of taking a closer look at the behavioral network a participant brings into the lab, and the behavioral network that emerges during the experiment. We also emerge a little more skeptical of mechanism advocates who have not looked at such data. On balance, we lean into the positive: behavioral deviation seems a rich source of insight into agency under social structural constraint, the juice that makes the Coleman boat go.

REFERENCES

Ahuja G (2000) Collaboration networks, structural holes, and innovation: A longitudinal study. *Adm. Sci. Quart.* 45(3):425-455.

Ali RF, Dominic PDD, Ali SEA, Rehman M, Sohail A (2021) Information security behavior and information security policy compliance: A systematic literature review for identifying the transformation process from noncompliance to compliance. *Appl. Sci.* 11:3383.

Angrist JD (2006) Instrumental variables methods in experimental criminological research: what, why, and how. *J. Exp. Criminology* 2:23-44.

Badura KL, Grijalva E, Newman DA, Yan TT, Jeon G (2018) Gender and leadership emergence: A meta-analysis and explanatory model. *Pers. Psych.* 71(3):335-367.

Bales RF (1950) *Interaction Process Analysis*. Chicago, IL: University of Chicago Press.

Bothner MS, Kang JH, Stuart TE (2007) Competitive crowding and risk taking in a tournament: Evidence from NASCAR racing. *Admin.Sci. Quart.* 52(2):208-247.

Brands RA, Mehra A (2019) Gender, brokerage, and performance: A construal approach. *Acad. Manag. J.* 62(1):196-219.

Brands, RA, Kilduff M (2014) Just like a woman? Effects of gender-biased perceptions of friendship network brokerage on attributions and performance. *Org. Sci.* 25(5):1530-1548.

Brands, RA, Ertug G, Fonti F, Tasselli S (2022) Theorizing gender in social network research: What we do and what we can do differently. *Acad. Manag. Ann.* 16(2):588-620.

Brass DJ (2022) New developments in social network analysis. *Ann. Rev. Organ. Psych. and Organ. Beh.* 9:8.1-8.22.

Burger MJ, Buskens V (2009) Social context and network formation: An experimental study. *Soc. Networks* 31: 53-75.

Burt RS (1984) Network items and the General Social Survey. *Soc. Networks* 6:293-339.

Burt RS (1987) Social contagion and innovation: Cohesion versus structural equivalence. *Am. J. Sociol.* 92(6):1287-1335.

Burt RS (1992) *Structural Holes*. Cambridge, MA: Harvard University Press.

Burt RS (1998) The gender of social capital. *Rationality and Society* 10(1):5-46.

Burt RS (2002) Bridge decay. *Soc. Networks* 24:333-363.

Burt RS (2004) Structural holes and good ideas. *Am. J. Soc.* 110(2):349-399.

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

Burt RS (2010) *Neighbor Networks*. New York: Oxford University Press.

Burt RS (2012) Network-related personality and the agency question: Multirole evidence from a virtual world. *Am. J. Sociol.* 118(3):543-91.

Burt RS (2021) Structural holes capstone, cautions, and enthusiasms. Pp. 384-416 in Small ML, Perry BL, Pescosolido B, & Smith E (Eds.), *Personal Networks*. New York: Cambridge University Press.

Burt RS and Carlton DS (1989) Another look at the network boundaries of American markets. *Am. J. Sociol.* 95(3):723-753.

Burt RS, Merluzzi J (2016) Network oscillation. *Acad. Manag. Disc.* 2(4): 368-391.

Burt RS, Opper S (2024) *Guanxi* and structural holes: Strong bridges from relational embedding. *Am. J. Sociol.* 130(1):1-43.

Burt RS, Opper S (2026) *Strong Bridges. Trust Beyond Structure*. New York: Oxford University Press.

Burt RS, Opper S, Holm HJ (2022) Cooperation beyond the network. *Org. Sci.* 33(2):495-517.

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

Burt RS, Reagans RE, Volvovsky HC (2021) Network brokerage and the perception of leadership. *Soc. Networks* 65:33-50.

Burt RS, Talmud I (1993) Market niche. *Soc. Networks* 15(2):133-149.

Cao J, Smith EB (2021) Why do high-status people have larger social networks? Belief in status-quality coupling as a driver of network-broadening behavior and social network size. *Org. Sci.* 32(1):111-132.

Carnabuci G, Emery C, Brinberg D (2018). Emergent leadership structures in informal groups: A dynamic, cognitively informed network model. Org. Sci. 29(1):118-133.

Carnabuci G, Quintane E (2023) When people build networks that hurt their performance: Structural holes, cognitive style, and the unintended consequences of person-network fit. *Acad. Manag. J.* 66(5):1360-1383.

Chase ID (1980) Social process and hierarchy formation in small groups: A comparative perspective. *Am. Sociol. Rev.* 45(6):905-924.

Christie LS, Luce RD, Macy J Jr. (1952) Communication and learning in task-oriented groups. Technical Report No. 231, Research Laboratory of Electronics, MIT (RLE-TR-231-00040063.pdf).

Coleman JS (1990) *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

DiMatteo MR (2004) Variations in patient's adherence to medical recommendations: A quantitative review of 50 years of research. *Medical Care* 42(3):200-209.

Emirbayer M, Mische A (1998) What is agency? *Am. J. Sociol.* 103(4):962-1023.

Fang R, Landis B, Zhang Z, Anderson MH, Shaw JD, Kilduff M (2015) Integrating personality and social networks: A meta-analysis of personality, network position, and work outcomes in organizations. *Org. Sci.* 26(4):1243-1260.

Fang R, Zhang Z, Shaw JD (2021) Gender and social network brokerage: A meta-analysis and field investigation. *J. Appl. Psych.* 106(11):1630-1654.

Freeman LC (1979) Centrality in social networks: II. Experimental results. *Soc. Networks* 2(2):119-161.

Freeman LC (2004) *The Development of Social Network Analysis*. Vancouver, Canada: Empirical Press.

Gould RV (2002) The origins of status hierarchies: A formal theory and empirical test. *Am. J. Sociol.* 107(5):1143-1178.

Gould RV (2003) *Collision of Wills*. Chicago, IL: University of Chicago Press.

Greenberg, J (2021) Social network positions, peer effects, and evaluation updating: An experimental test in the entrepreneurial context. *Org. Sci.* 32(5): 1174-1192.

Gulati R, Srivastava SB (2014) Bringing agency back into network research: Constrained agency and network action. *Res. Sociol. Organ.* 40: 73-93.

Halevy N, Halali E, Zlatev JJ (2019) Brokerage and brokering: An integrative review and organizing framework for third party influence. *Acad. Manag. Ann.* 13 (1): 215-239.

Hardin G (1968) The tragedy of the commons. *Science* 162(3859):1243-1248.

Hedstrom P, Swedberg R (1998) *Social Mechanisms*. New York: Cambridge University Press.

Hummon NP, Doreian P, Freeman LC (1990) Analyzing the structure of the centrality-productivity literature created between 1948 and 1979. *Sci. Comm.* 11:459-480.

Jancsics D, Espinosa S, Carlos J (2023) Organizational noncompliance: An interdisciplinary review of social and organizational factors. *Management Rev. Quart.* 73:1273-1301.

Janicik GA, Larrick RP (2005) Social network schemas and the learning of incomplete networks. *J. Pers. Soc. Psych.* 88(2):348-364.

Kilduff GJ (2014) Driven to win: Rivalry, motivation, and performance. *Soc. Psych. Personality Sci.* 5(8): 944-952.

Kilduff GJ, Galinsky AD, Gallo E, Reade JJ (2016) Whatever it takes to win: Rivalry increases unethical behavior. *Acad. Manag. J.* 59(5):1508-1534.

Kilduff M, Wang K, Lee SY, Tsai W, Chuang YT, Tsai FS. (2024) Hiding and seeking knowledge providing ties from rivals: A strategic perspective on network perceptions. *Acad. Manag. J.* 67(5):1207-1233.

Kleinbaum AM, Stuart TE (2014) Inside the black box of the corporate staff: Social networks and the implementation of corporate strategy. *Strat. Manag. J.* 35(1):24-47.

Koka BR, Prescott JE (2008) Designing alliance networks: The influence of network position, environmental change, and strategy on firm performance. *Strat. Manag. J.* 29(6): 639-661.

Kwon SW, Rondi E, Levin DZ, De Massis A, Brass DJ (2020) Network brokerage: An integrative review and future research agenda. *J. Manag.* 46(6):1092-1120.

Leavitt HJ (1951) Some effects of certain communication patterns upon group performance. *J. Abnormal Psych.* 46: 38-50.

Leavitt HJ (1996) The old days, hot groups, and manager's lib. *Admin. Sci. Quart.* 41:288-300.

Little RJ, Yau LHY (1998) Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model. *Psych. Methods* 3(2):147-159.

Marsden PV (1987) Core discussion networks of Americans. *Am. Sociol. Rev.* 52(1): 122-131.

Marsden PV, Fekete M, Baum DS (2021) Egocentric network studies within the General Social Survey: Measurement methods, substantive findings, and methodological research. Pp. 519-551 in Small ML, Perry BL, Pescosolido BA, Smith EB (Eds.), *Personal Networks*. New York: Cambridge University Press.

Martin JL (2011) *The Explanation of Social Action*. New York: Oxford University Press.

Martin JL, Yeung KT (2006) Persistence of close personal ties over a 12-year period. *Soc. Networks* 28:331-362.

Martin JL, Overgoor J, State B (2024) Persistence and change in structural signatures of tie formation over time. *Soc. Networks*. 77:5-16.

McClellan M, McNeil BJ, Newhouse JP (1994) Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *J. Am. Med. Ass.* 272(11):859-866.

Mehra A, Kilduff M, Brass DJ (1998) At the margins: A distinctiveness approach to the social identity and social network of underrepresented groups. *Acad.Management J.* 41(4):441-52.

Mehra A, Kilduff M, Brass DJ (2001) The social networks of high and low self-monitors: Implications for workplace performance. *Admin. Sci. Quar.* 46(1):121-146

Meyer RD, Kelly ED, Bowling NA (2018). Situational strength theory: A formal conceptualization of a popular idea. Pp. 79-95 in Rauthmann JF, Sherman RA, Funder DC (Eds), *The Oxford Handbook of Psychological Situations*. Oxford: Oxford University Press.

Mischel W (1968) *Personality and Assessment*. New York: Wiley.

Mizruchi MS, Mariolis P, Schwartz M, Mintz B (1986) Techniques for disaggregating centrality scores in social networks. *Sociol. Methodol.* 16:26-48.

Olsen M (1965) *The Logic of Collective Action*. Cambridge, MA: Harvard University Perss.

Orosz K, Proteasa V, Craciun D (2021) The use of instrumental variables in higher education research. *Theory Method High. Educ. Res.* 6:61-80.

Piezunka H, Lee W, Haynes R, Bothner MS (2018) Escalation of competition into conflict in competitive networks of Formula One drivers. *Proc. Natl. Acad. Sci.* 115(15):E3361-E3367.

Polan S, Schitter C (2018) Prolific.ac — A subject pool for online experiments. *J. Behav. Exp. Finance* 17:22-27.

Podolny JM (1993) A status-based model of market competition. *Am. J. Sociol.* 98(4):829-872.

Podolny JM (2005) *Status Signals*. Princeton, NJ: Princeton University Press.

Porac JF, Thomas H, Wilson F, Paton D, Kanfer A (1995) Rivalry and the industry model of Scottish knitwear producers. *Admin. Sci. Quart.* 40(2):203-227.

Reagans RE, Burt RS (2026) Assessing the extent to which network structure causes people to be perceived as a leader. *Network Sci.*, In Press.

Reagans RE, Burt RS, Liu DD (2026) Jargonization, language development, and team performance. *Sociol. Sci.* 13:314-331.

Ridgeway CL (2001) Gender, Status, and Leadership. *J. Soc. Issues* 57(4): 637-655.

Sagarin BJ, West SG, Ratnikov A, Homan WK, Ritchie TD, Hansen EJ (2014) Treatment noncompliance in randomized experiments: Statistical approaches and design issues. *Psych. Methods* 19(3):317-333.

Sewell WH Jr. (1992). A theory of structure: Duality, agency, and transformation. *Am. J. Sociol.* 98(1):1-29.

Stube MJ (2005) What did Triplett really find? A contemporary analysis of the first experiment in social psychology. *Am. J. Psych.* 118:271-286.

Tasselli S, Kilduff M (2021) Network agency. *Acad. Manag. Ann.* 15(1):68-110.

Triplett N (1898) The dynamogenic factors in pacemaking and competition. *Am. J. Psych.* 9:507-533.

van Burg E, Elfring T, Cornelissen JP (2022). Connecting content and structure: A review of mechanisms in entrepreneur's social networks. *Int. J. Manag. Rev.* 24(2):188-209.

Vedres B, Vasarhelyi O (2019) Gendered behavior as a disadvantage in open source software development. *EPJ Data Science* 8(25): doi.org/10.1140/epjds/s13688-019-0202-z.

Vissa B (2012) Agency in action: Entrepreneurs' networking style and initiation of economic exchange. *Org. Sci.* 23(2):492-510.

Webb EJ, Campbell DT, Schwartz RD, Sechrest L (1966) *Unobtrusive Measures*. Chicago: Rand-McNally.

Zolnierek KBH, DiMatteo MR (2009) Physician communication and patient adherence to treatment. *Med. Care* 47(8):826-834.

Table 1. Team Learning Curves Are Unaffected by Gendered-Broker Emergence

	Duration	Number of Messages	Accuracy
Trial (ln)	-3.11*** (.42)	-.43*** (.05)	.28** (.09)
Maximum behavioral deviation toward brokerage (most negative behavioral - assigned constraint, e.g., -5.0 in Fig. 5)	-.05 (.04)	-.01 (.01)	.003 (.01)
Interaction, above times (ln(trial) – mean ln(trial))	-.07 (.06)	-.01 (.01)	.01 (.01)
Maximum behavior deviation toward brokerage is by a male (e.g., Figure 5)	-.70 (.56)	-.02 (.09)	.14 (.11)
Intercept	9.39	4.85	.81
R^2 or Pseudo R^2	.49	.27	.03
Test statistic and probability that the learning curve is independent of the three behavioral deviation predictors ($F_{(3, 222)}$ or χ^2 with 3 *df*)	1.98, $p = .12$	3.72, $p = .29$	2.80, $p = .42$

Note: Estimates are across 227 trials by teams assigned to CLIQUE treatment networks. First column contains OLS estimates predicting minutes per trial. Second and third columns contain Poisson estimates respectively predicting number of messages per trial and number of accurate solutions per trial. Test statistics in parentheses are estimated using Stata "vce(robust)" option (similar results using the "cluster" option to correct for autocorrelation between trials by the same team). *** $p < .001$ ** $p < .01$

Table 2. Connected Brokers Compete for Leadership

	Predict Behavioral Deviation by Paired Brokers		
	Coef.	S.E.	*t*
Lead Broker	-1.07	.42	-2.55
CB Network	4.95	2.19	2.26
Lead x CB	-15.52	2.65	-5.83

NOTE: Results are from OLS regression for graphs to the left, with robust estimates of standard errors (94 participants, 2.07 intercept, .48 R^2). Knowing gender does not improve the prediction (t = 0.40, p = .69), and lead brokers tending to be male fails to reject the null (49% vs. 39%; z = 0.60, p = .56).

Table 3. Team Learning Curves Are Unaffected by Broker Rivalry

	Duration	Number of Messages	Accuracy
Trial (ln)	-3.13*** (.25)	-.40*** (.04)	.20*** (.05)
CB network (versus DB network)	-.19 (.45)	-.04 (.08)	-.08 (.08)
Interaction, above times (ln(trial) – mean ln(trial))	-.21 (.54)	.01 (.08)	.11 (.12)
Behavioral deviation gap between the two brokers (network constraint on other broker minus constraint on lead broker)	.01 (.02)	-.002 (.004)	.002 (.004)
Interaction, above times (ln(trial) – mean ln(trial))	.003 (.03)	-.003 (.004)	-.007 (.005)
Intercept	9.81	4.92	1.08
R^2 or Pseudo R^2	.47	.29	.02
Test statistic and probability that the learning curve is independent of two behavioral deviation predictors ($F_{(4, 459)}$ or χ^2 with 4 *df*)	0.16, $p = .96$	2.88, $p = .58$	2.54, $p = .64$

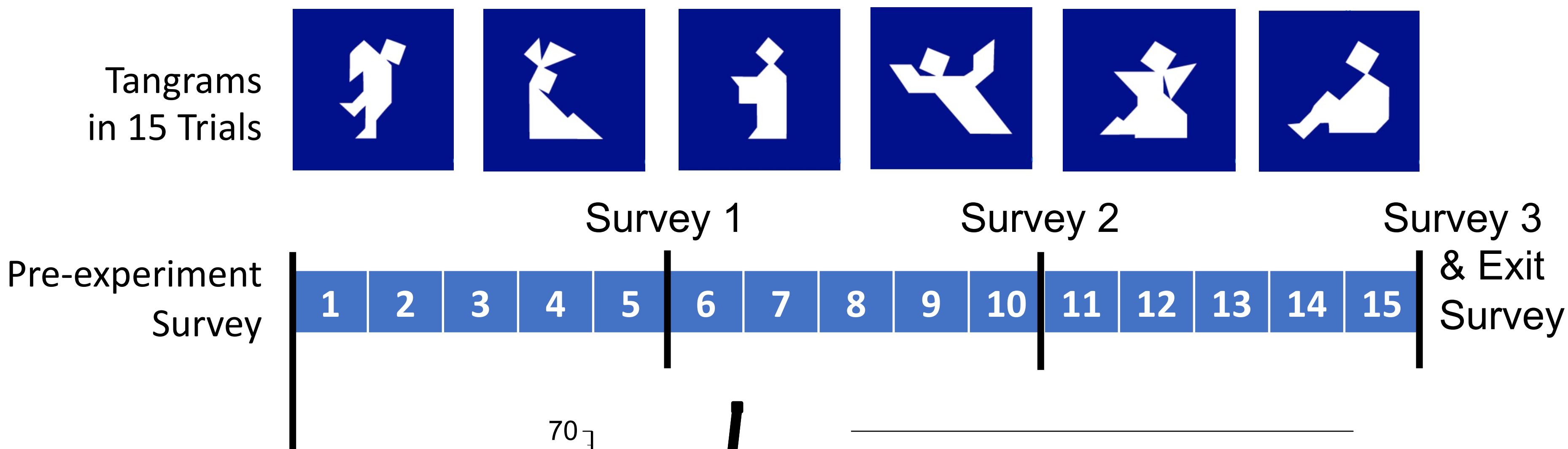
Note: Estimates are across 465 trials by teams assigned to CB or DB treatment networks. First column contains OLS estimates predicting minutes per trial. Second and third columns contain Poisson estimates respectively predicting number of messages per trial and number of accurate solutions per trial. Test statistics in parentheses are estimated using Stata “vce(robust)” option (similar results using ”cluster” option to correct for autocorrelation between trials by the same team). *** $p < .001$ ** $p < .01$

Table 4. Team Learning Curves Are Unaffected by Closed Networks Pre-Experiment

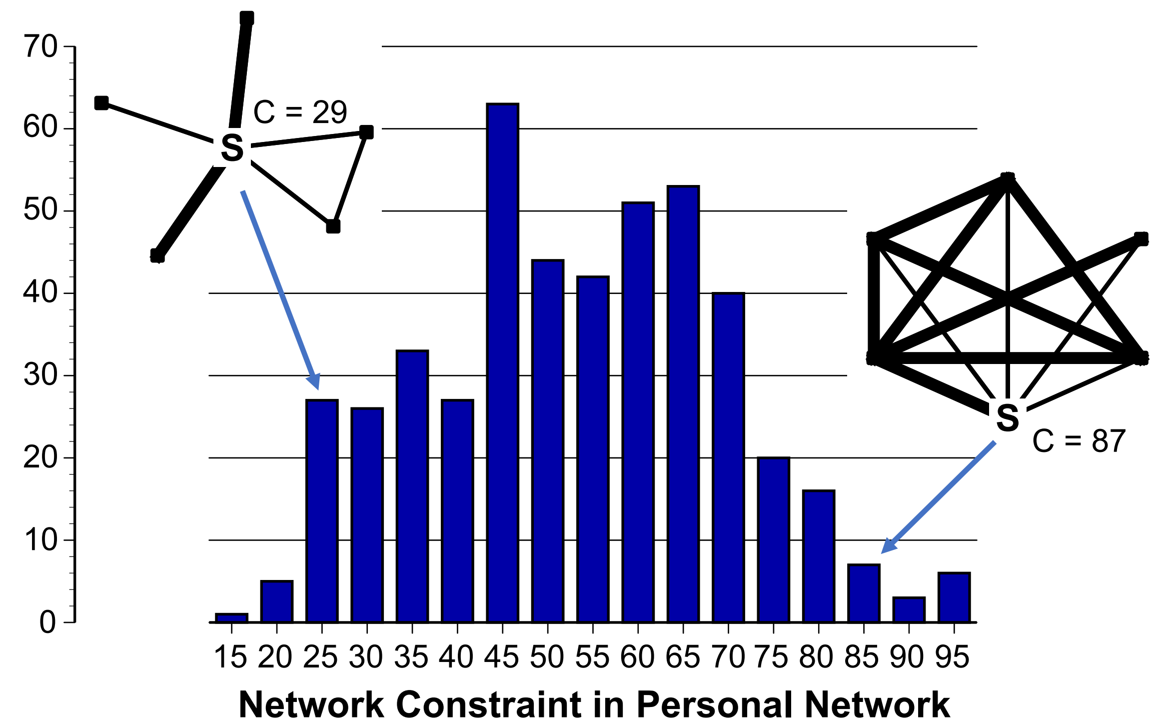
	Duration	Number of Messages	Accuracy
Trial (ln)	-3.08*** (.45)	-.33*** (.06)	.23* (.10)
Pre-experiment network constraint for teammate with <u>MOST</u> closed network	-.03 (.01)	-.001 (.002)	-.002 (.002)
Interaction, above times (ln(trial) – mean ln(trial))	.003 (.02)	.001 (.002)	.003 (.003)
Pre-experiment network constraint for teammate with <u>LEAST</u> closed network	.0003 (.01)	-.003 (.002)	.001 (.003)
Interaction, above times (ln(trial) – mean ln(trial))	-.003 (.02)	.004 (.0032)	.001 (.004)
Intercept	7.44	4.93	1.11
R^2 or Pseudo R^2	.48	.30	.03
Test statistic and probability that the learning curve is independent of two behavioral deviation predictors ($F_{(4, 426)}$ or χ^2 with 4 *df*)	1.66, *p* = .16	5.96, *p* = .20	1.39, *p* = .85

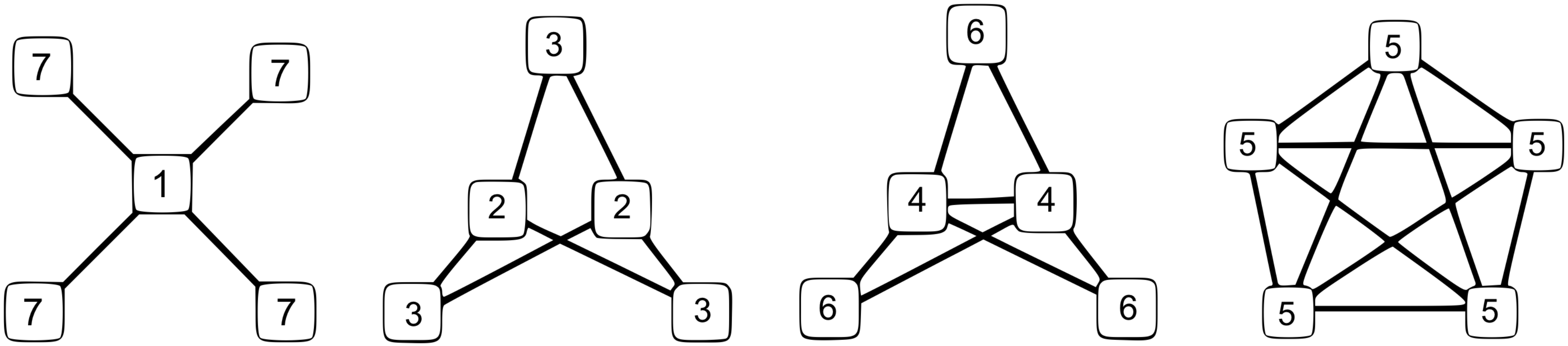
Note: Estimates are across 432 trials by teams assigned to either of the two treatment networks prone to behavioral deviation (CLIQUE and CB). First column contains OLS estimates predicting minutes per trial. Second and third columns contain Poisson estimates respectively predicting number of messages per trial and number of accurate solutions per trial. Test statistics in parentheses are estimated using Stata "vce(robust)" option (similar results using the "cluster" option to correct for autocorrelation between trials by the same team). *** $p < .001$ ** $p < .01$ * $p < .05$

Figure 1. Experiment Task & External Network



GSS name generator: "From time to time, most people discuss important matters with other people, people they trust. The range of important matters varies from person to person across work, leisure, family, politics, whatever. The range of relations varies across work, family, friends, and advisors. *If you look back over the last six months, who are the five people with whom you most discussed matters important to you?*





WHEEL Network

Mixed (DB) Network: Disconnected-Brokers

Mixed (CB) Network: Connected-Brokers

CLIQUE Network

Low-Constraint Positions

Position ID	Treatment Constraint	Structural Holes	S.D. Behave. Constraint
1	**25**	**6**	**4.39**
2	**33**	**3**	**1.54**

Position ID	Treatment Constraint	Structural Holes	S.D. Behave. Constraint
3	**50**	**1**	**1.23**
4	**68**	**3**	**12.16**

High-Constraint Positions

Position ID	Treatment Constraint	Structural Holes	S.D. Behave. Constraint
5	**77**	**0**	**7.31**
6	**100**	**0**	**8.75**
7	**100**	**0**	**0.00**

Figure 2.

Four Treatment Networks.

Figure 3. Participant-Machine Interface

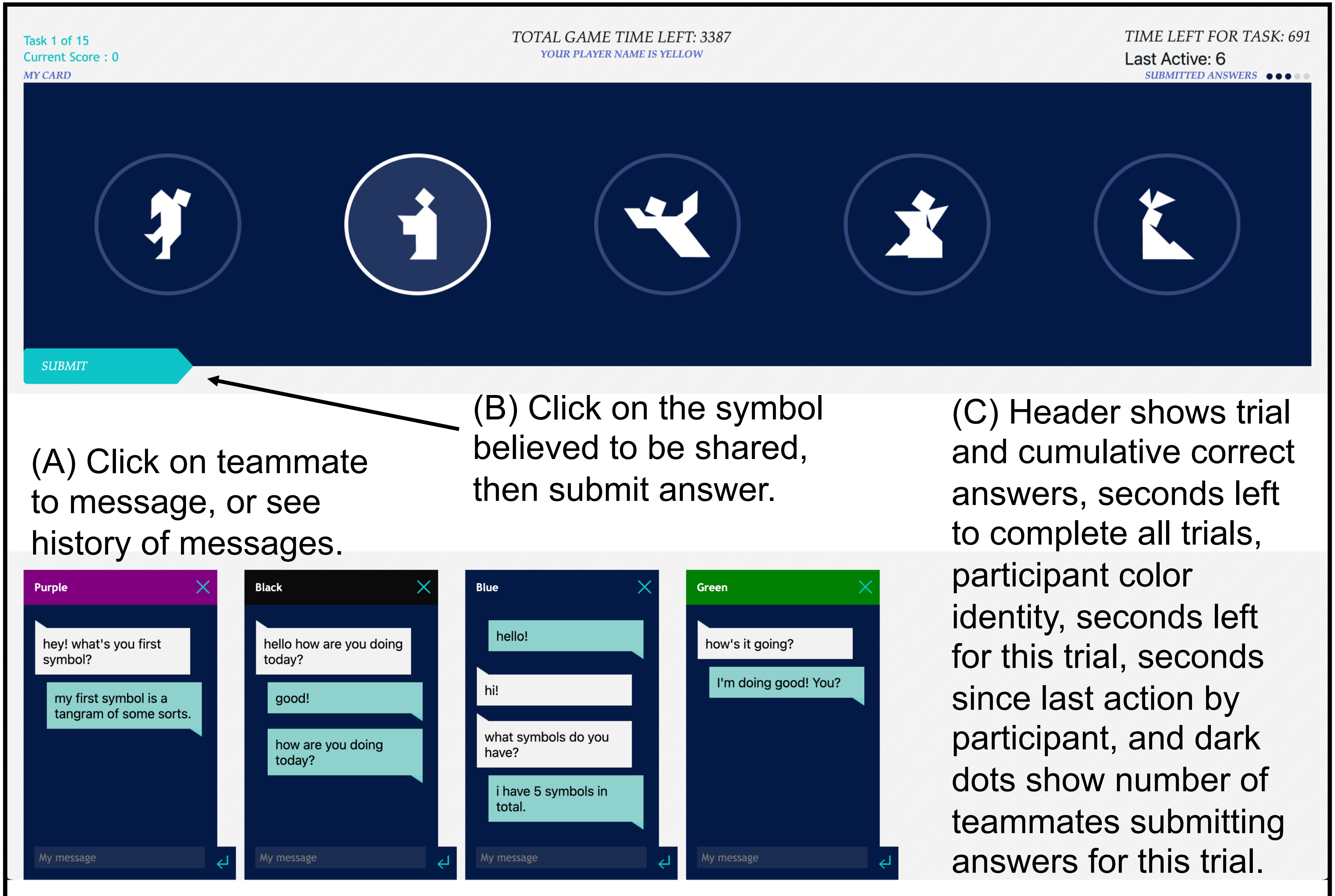
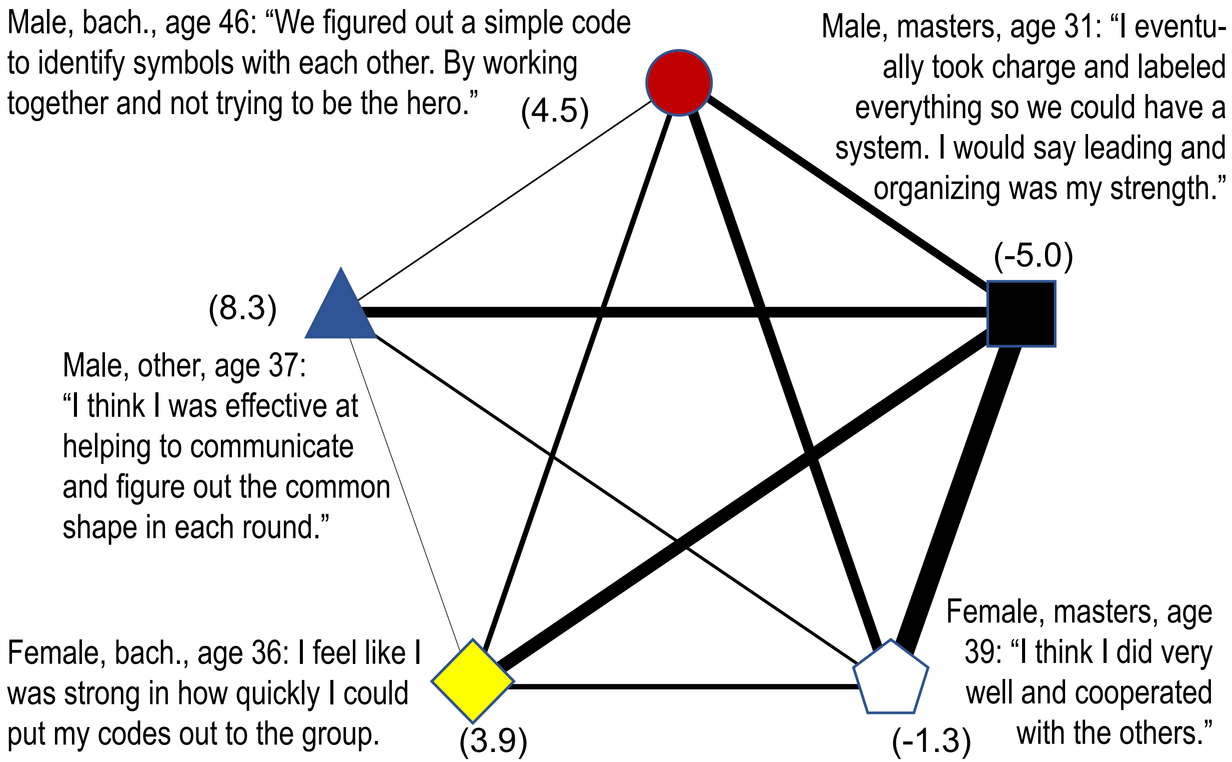


Figure 4.

Illustration, Gendered-Broker Hypothesis



NOTE: Behavioral deviation scores in parentheses (behavioral constraint - treatment constraint). Quotes are responses to exit question: "How would you describe your strength in the game?" No quotes on abbreviated responses. Square receives all but one of the team leader citations. (Color coding in experiment is augmented here by shapes to accommodate black and white printing.) Other attributes are gender, education, and age.

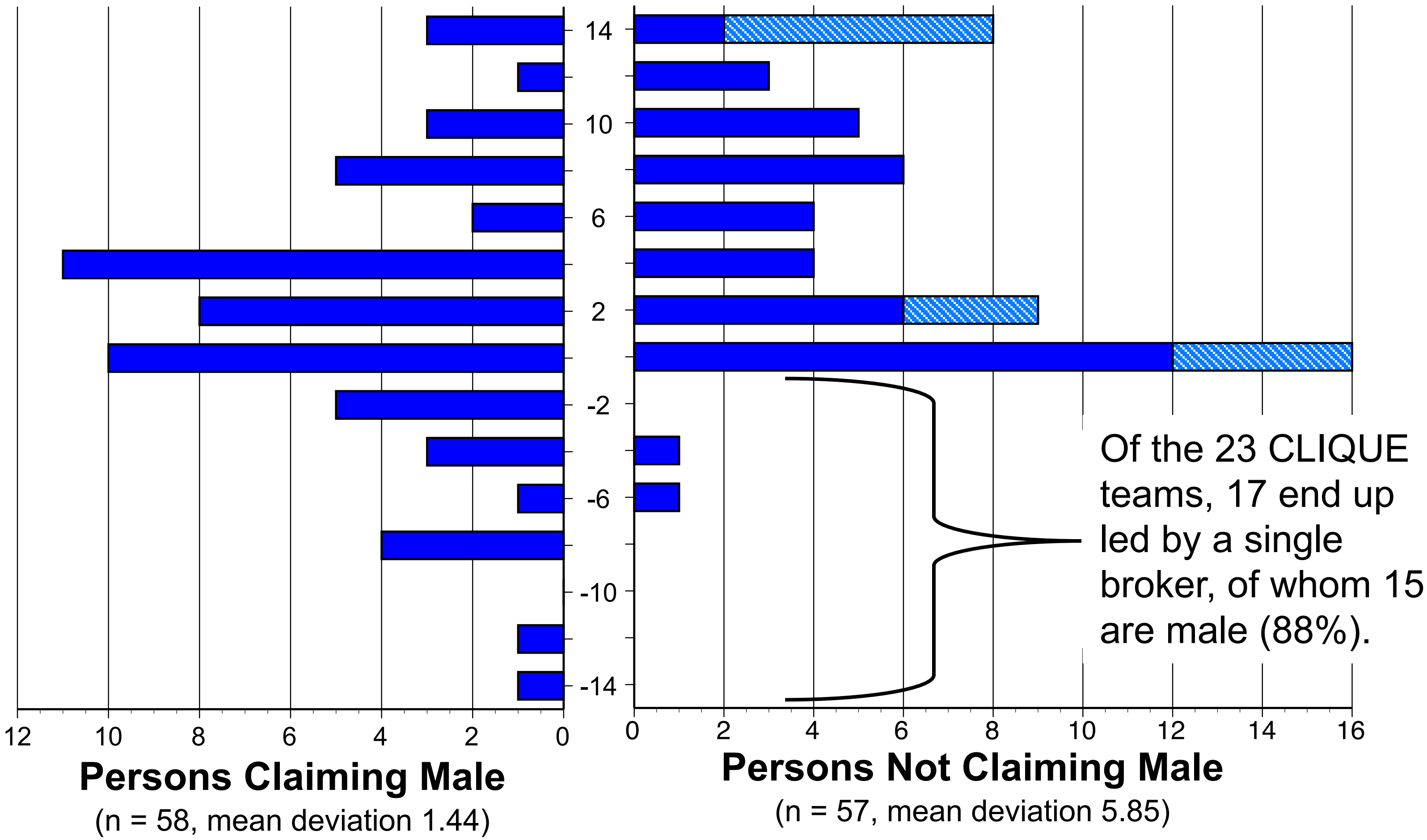
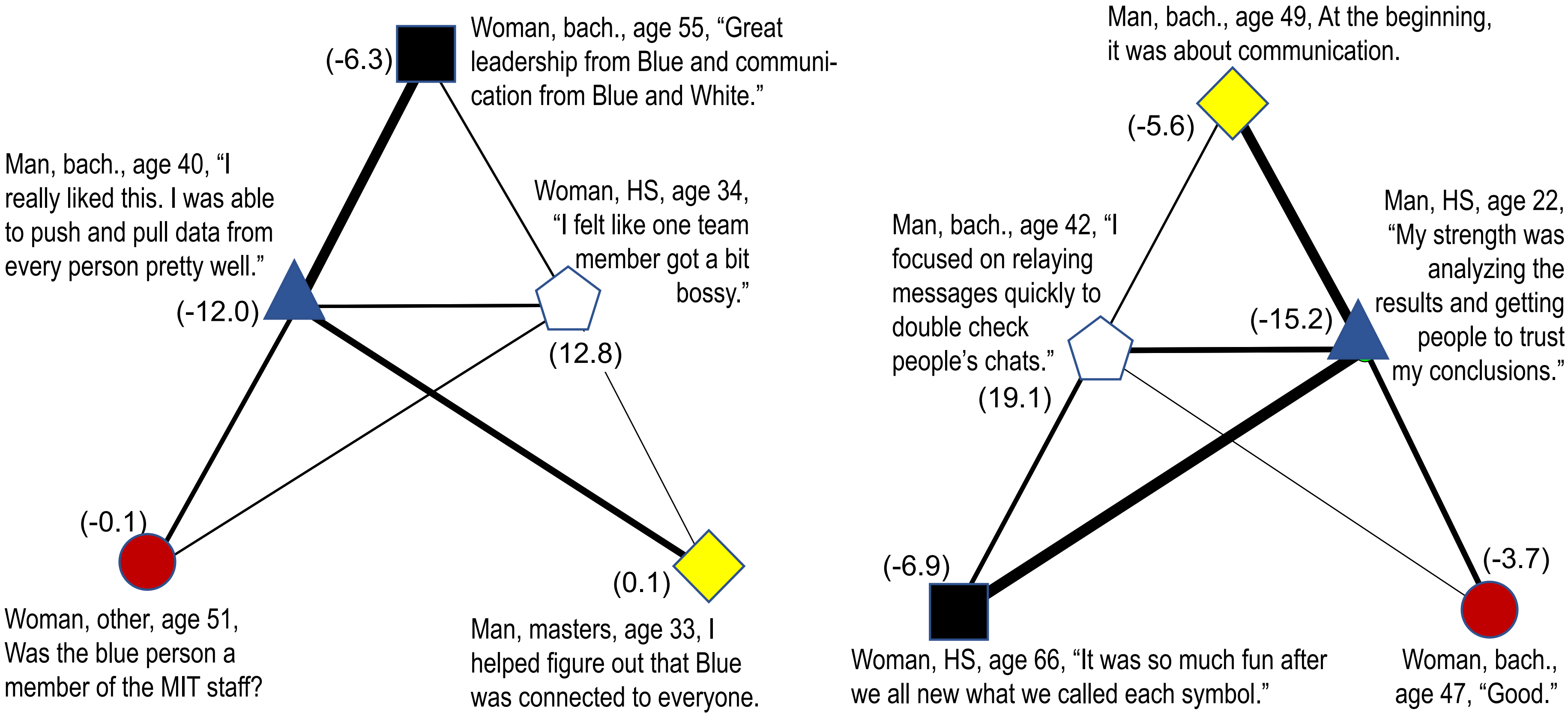


Figure 5. Gendered-Broker Hypothesis

NOTE: Vertical axis is constraint in behavioral network minus constraint in CLIQUE treatment networks (difference rounded down to nearest two-integer absolute value, with extremes truncated at ±14). Participants who did not claim a male or female gender are included in the hatched areas. The six teams in which a single broker did not emerge as leader typically had dual leadership shared between a greater and a lesser broker (1 male pair, 1 female pair, and 4 mixed gender pairs). In fact, not many teams assigned to CLIQUE treatment networks actually experienced a CLIQUE network.

Figure 6. Assigned to CB Network, These Teams Experience a WHEEL Network



A. Team 15 Behavioral Network

(Blue is lead broker and receives 100% of leader citations. Deviation gap between brokers is 13.2+11.6 = 24.8)

B. Team 66 Behavioral Network

(Green is lead broker and receives 100% of leader citations. Deviation gap between brokers is 19.5+14.8 = 34.3)

NOTE: Thicker lines indicate more messages. Behavioral deviation score is in parentheses (behavioral – treatment constraint). Quotes are response to question: “How would you describe your strength in the game?” No quotes on abbreviated responses. Other attributes are gender, education, and age. Color coding in experiment is augmented here by shapes for black and white printing.

Figure 7. Connected Brokers Compete for Leadership

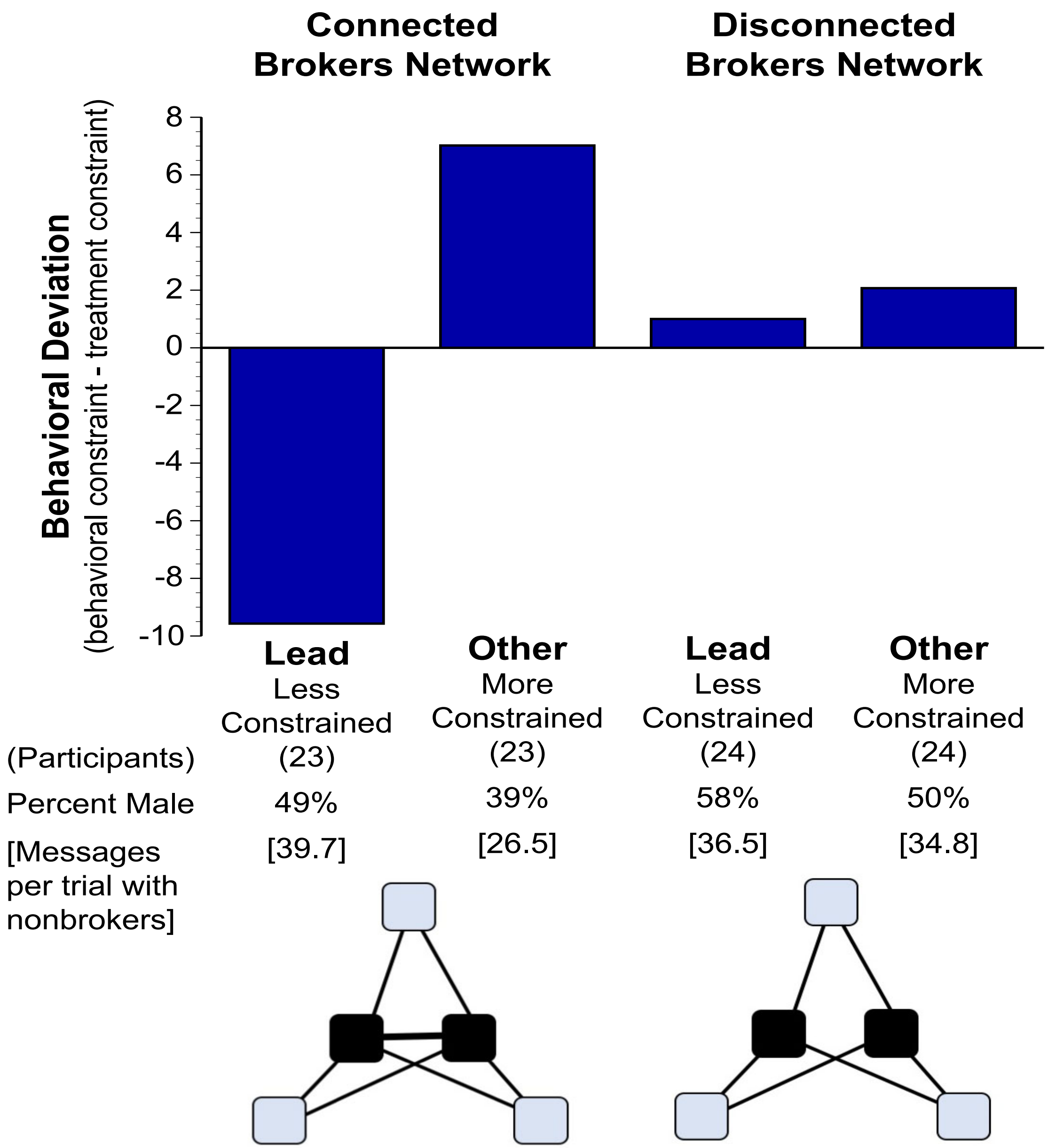
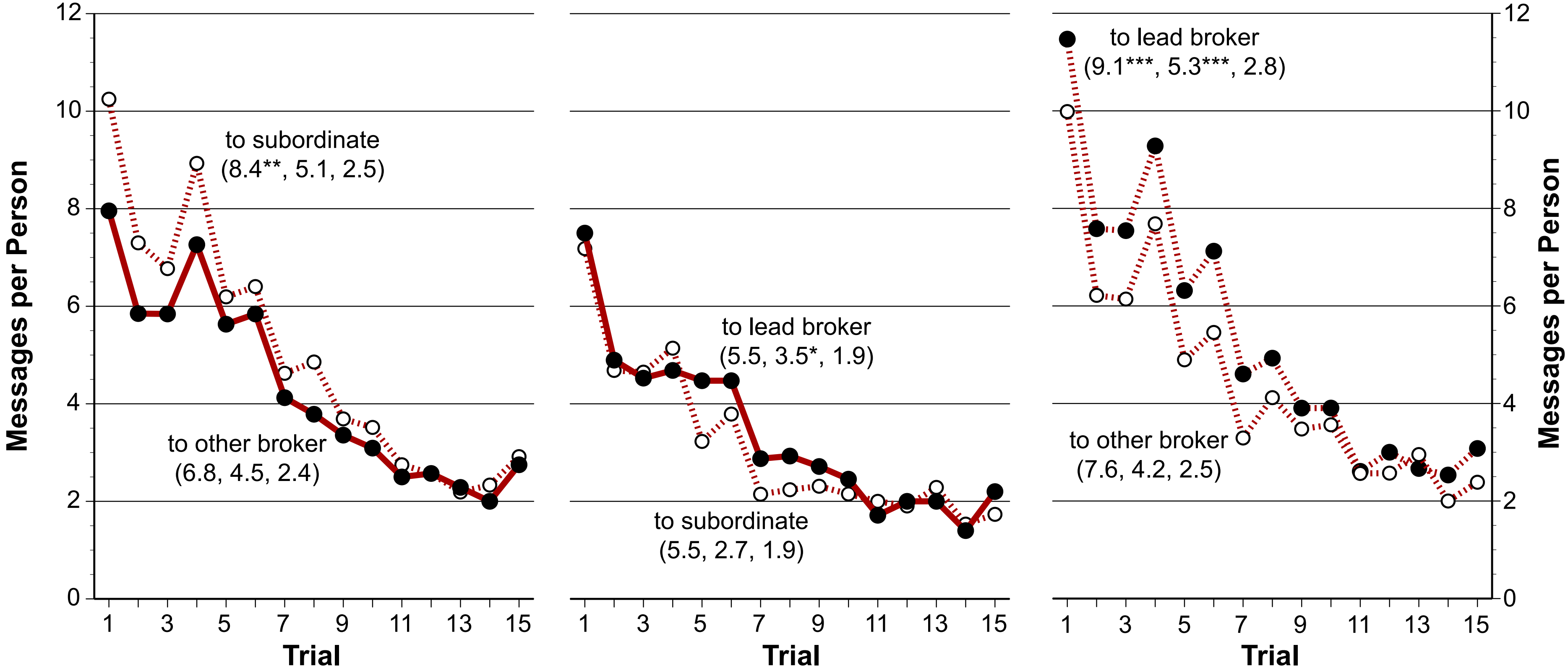


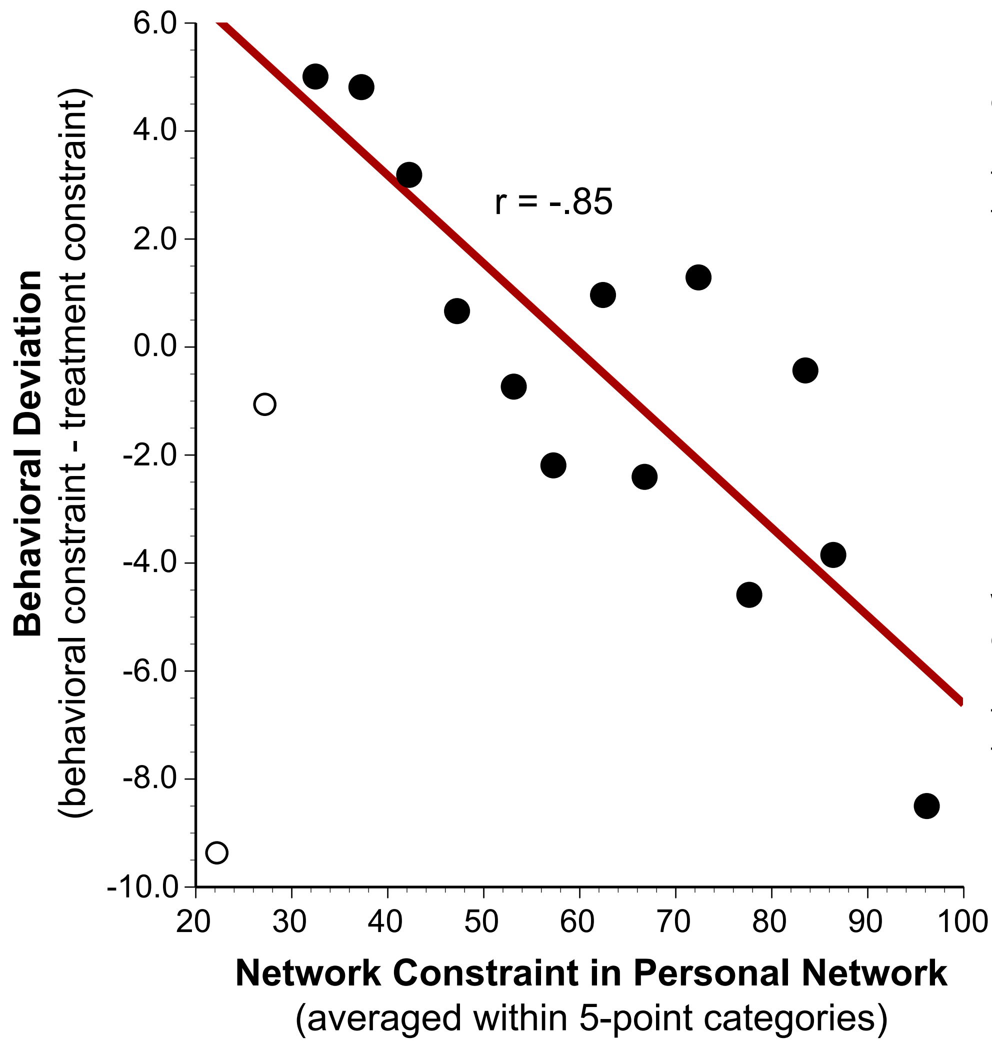
Figure 8.
Lead Broker Is Established Early via Constituents



A. Messages from Lead Broker | **B. Messages from Other Broker** | **C. Messages from Subordinates to Brokers**

NOTE: For teams assigned to a connected-brokers treatment network (Figure 4), these are learning curves showing the mean number of messages sent in each trial to each type of teammate. For example, graph A shows a mean of 8.0 messages sent by lead brokers to the other broker during the first trial, and a mean of 10.3 messages to each of the three subordinate teammates. Numbers in parentheses are mean messages sent in novice, intermediate, and mature trials (respectively 1-4, 5-10, and 11-15). Asterisks indicate message counts higher than the corresponding count in the same graph (e.g., 8.4 is higher than 6.8 at a .01 level of confidence). ***$p \leq .001$ **$p \leq .01$ *$p \leq .05$

Figure 9. More Closed Outside, More Open Inside



Up here are people who behave during the experiment so as to end up more constrained than was assigned by treatment (e.g., triangle in Figure 4, pentagons in Figure 6).

Down here are people who behave during the experiment so as to end up less constrained than was assigned by treatment (e.g., square in Figure 4, triangles in Figure 6).

NOTE: Graph only includes participants in the two treatment networks that facilitate deviation (CB & CLIQUE). Averages within five-point intervals of constraint in personal network are plotted. Correlation is through solid dots (exclude outliers, hollow dots).

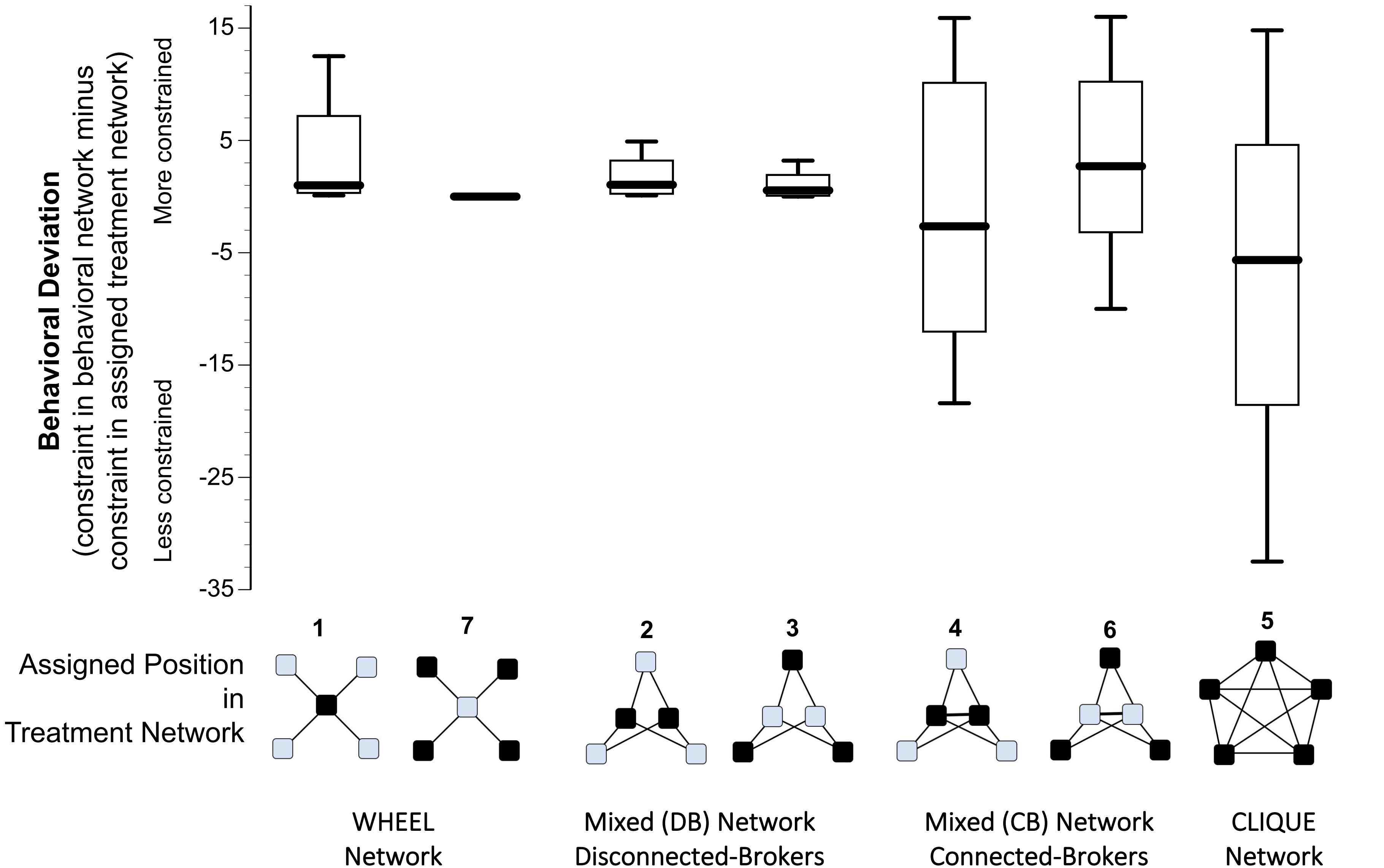


Figure 10. Behavioral Deviation by Treatment Network

Percentiles 25, 50, 75 define box. Whiskers to 5th and 10th percentiles.