A Novel Approach to Studying the Role Influence Plays in Team Collective Intelligence

Lisa R. O'Bryan[1,2], Timothy Oxendahl[1], Simon Garnier[3], Santiago Segarra[2], Matthew Wettergreen[4], Ashutosh Sabharwal[2], Margaret E. Beier[1]

ABSTRACT

Studying collective intelligence in teams poses key challenges because individual contributions to team outcomes – such as ideas – can be abstract and difficult to measure. Although many studies have examined how team members shape collective outputs, they often utilize indirect measures of influence that focus on perceptions, interactions, and performance, which can differ from actual influence. In this study, we develop an objective measure of influence within a creative design task. We operationalized influence as the objective similarity between designs created in an independent Brainstorming Phase and subsequent collaborative (or independent) Design Phase. We demonstrate the utility of our approach by evaluating variation in influence across team members, examining individual characteristics correlated with influence, and analyzing how the characteristics of influential team members correlate with team performance. Key takeaways are that teams gravitated towards some ideas more than others, especially the ideas of team members who had greater expertise, who made more work contributions, and whose ideas were more similar to other team members' ideas. Our approach demonstrates the utility of our objective influence measure in enabling a mechanistic understanding of idea integration within teams, and we discuss how this approach can foster new insights within the field of collective intelligence.

---

[1]Department of Psychological Sciences, Rice University, Houston, TX, USA
[2]Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA
[3]Department of Biological Sciences, New Jersey Institute of Technology, Newark, NJ, USA
[4]Department of Bioengineering, Rice University, Houston, TX, USA

Author Note:
Funding for this project was provided by a Microsoft Productivity Research Grant and the National Science Foundation (Award Number: 1910117).
The authors have no conflicts of interest to disclose.
Correspondence concerning this manuscript should be addressed to Dr. Lisa O'Bryan, Rice University, Department of Electrical and Computer Engineering, MS-366, 6100 Main Street, Houston, TX 77005. Email: obryan@rice.edu, Phone: 713-385-9257

Significance Statement

Teams are increasingly prevalent within the workplace, serving as a key organizational strategy for enhancing problem-solving, innovation, and performance. However, there is limited understanding of how individual ideas are (and should be) integrated into team outcomes during open-ended problem-solving tasks. Our study demonstrates how the use of an objective measure of influence can provide insight into this topic. Utilizing an online design game, we developed an objective measure of the influence individuals' initial ideas had over subsequent designs when either working collaboratively or independently. By assessing measurable design changes over time we avoid pitfalls associated with relying upon more subjective or indirect measures of influence, which may not accurately reflect how team member inputs shape team outcomes. Results suggest that teams can enhance their performance by ensuring that those who are higher in expertise are free to influence team outcomes. Overall, our study highlights the utility of objective measures of influence in clarifying the behavioral mechanisms underlying idea integration, individual influence, and team performance, and we discuss how these approaches can be leveraged to test critical hypotheses related to the emergence of collective intelligence.

1. Introduction

The enhanced performance of groups relative to individuals is a critical emergent property of collective behavior. This phenomenon has been documented in diverse contexts, ranging from animal groups (Seeley & Buhrman, 1999) to human organizations (Riedl et al., 2021) to artificial swarms (Bonabeau et al., 1999). When it comes to human teams, the enhanced performance of teams over individuals is frequently termed "collective intelligence" (O'Bryan et al., 2020). Within teams, individual contributions to team outcomes often consist of knowledge, information, or ideas, which can combine with an individual's behaviors and traits to influence team outcomes (Bottger, 1984; Sherf et al., 2018). A central question in the study of collective intelligence is: How do the contributions of individual group members influence the emergence of intelligent group outcomes (De Dreu et al., 2008)?

Studies of collective decision-making and intelligence in teams often include proxy measures of individual influence over team outcomes (March, 1955). The three most common categories of influence measures are attributed influence (i.e., self-report), measures of interaction, and shifts in opinion or behavior (March, 1955). For example, to examine how team members influence team outcomes, studies may assess (1) team member perceptions, which are self-report measurements of constructs such as influence, voice, and expertise utilization (*attributed influence*; Bottger, 1984; Sherf et al., 2018), (2) key behaviors such as speaking time and work contributions (*measures of interaction*; Engel et al., 2014; Mast, 2002; Woolley et al., 2010), or (3) comparisons of outcomes or performance across an individual and team phase (*shifts in opinion* *or behavior*; Bottger, 1984; Goldman, 1965). Although such research has provided insight into team decision-making and collective intelligence, these measures of influence can greatly differ from one another (March, 1955, 1956), impeding comparison across studies. In addition, as some measures of influence may differ from more objective measures that take into account individuals' actual impacts on group outcomes (Bottger, 1984; March, 1956), they can cloud the understanding of how individual contributions shape group outcomes. In our study, we develop an objective measure of influence that expands upon measures of influence based on *shifts in opinion or behavior* within the context of an open-ended design task. We demonstrate how this approach can clarify the behavioral mechanisms underlying the emergence of collective intelligence in teams and discuss its potential to provide new insights into team effectiveness.

1.1 Measures of influence

Self-report surveys provide a relatively simple way for researchers to measure attributed influence in teams. For example, in their study of social influence within teams, Deuling (2011) assessed perceptions of team members' influence at multiple time points across the academic year and found that cognitive ability and extraversion were correlated with attributed influence. In addition to their simplicity, self-report measures of attributed influence also have the potential to capture features of interactions that may be more elusive or less readily measured by more objective measures, such as internal cognitions and attitudes. However, measures of attributed influence are subject to the same limitations as any self-report measure, such as bias, inconsistency, the limitations of memory, and invasiveness (Kihlstrom et al., 1999; Kozlowski & Chao, 2018). Furthermore, perceptions of influence may not correlate with more objective measures of influence (Bottger, 1984; March, 1956), calling into question whether self-report measures of attributed influence accurately reflect consequential influence processes within teams.

Measures of interaction focus on capturing influence attempts or proxies of influence attempts, which often take the form of communication acts, such as the production of speaking turns. Indeed, the quantity of communication alone is commonly associated with perceptions of influence, expertise, and leadership (Bales, 1953; Riecken, 1958). For example, Riecken (1958) found that ranked perceptions of team members' speaking times positively correlated with ranked perceptions of their contributions to the team outcome. However, March (1955) identifies three challenges with measures of influence that target interactions. Specifically, units of interaction may have different impacts on influence based upon 1) who they are directed from, 2) who they are directed to, and 3) their content. Furthermore, these measures do not take into account any changes in opinion or behavior that may occur as a result of these interactions. In fact, previous studies have found that interaction measures may not correlate with more objective measures of influence (Bottger, 1984). However, some studies have found that perceived expertise may mediate the relationship between measures of communication quantity and influence (Littlepage et al., 1995).

Measuring shifts in opinion requires measuring opinion or behavior during at least two time points and recording the change that occurs. In this approach, there is an assumption that the changes that one observes are due to interaction processes taking place between time points;

interactions that may impact the level of influence that one individual has on another or on their team. Although this approach can involve self-report measures of opinion, it differs from the self-report assessment of attributed influence described earlier in that researchers are not interested in the absolute value of the self-report, but rather the change from t_n to t_{n+1}. In other words, it is the change in opinion or behavior that suggests the presence of influence. Beyond self-report assessments of opinion, this approach may also leverage more objective measures, such as individual or team outputs. For example, within the context of collective intelligence, Bernstein et al. (2018) examined how individuals' solutions to a challenge changed over time after being exposed to others' ideas. With regards to team decision-making, some studies have compared performance between t_0 and t_1, with individuals working independently at t_0 and working together at t_1 (Goldman, 1965). This approach can highlight performance changes that occur when individuals with different levels of expertise work together. However, a drawback of this approach is that features of the individual and team outputs (e.g., their design or composition) are not considered, preventing examination of whether or how team members' ideas influenced the team output.

The studies that have most directly tested the shifts in opinion or behavior that occur when team members work together are those that directly measure changes in the features of outputs between an individual (t_0) and a team (t_1) phase. Most studies that have used this approach have utilized tasks that involve ranking items in a list, including the NASA moon landing task (Bottger, 1984) the Dessert Survival task (Littlepage et al., 1995), and other experimenter-derived rankings (March, 1956). As part of these tasks, researchers compare how closely individuals' initial rankings correspond with their final group ranking to calculate an objective measure of the level of influence each individual had over the group outcome. For some tasks, individual and group rankings can also be compared to the rankings provided by an expert to obtain an objective measure of ranking quality (Bottger, 1984; Littlepage et al., 1995). The benefits of this approach are that it is possible to directly link features of individual ideas to features of group outcomes, which can facilitate a more precise understanding of how individuals influence team outcomes (Bottger, 1984; March, 1956).

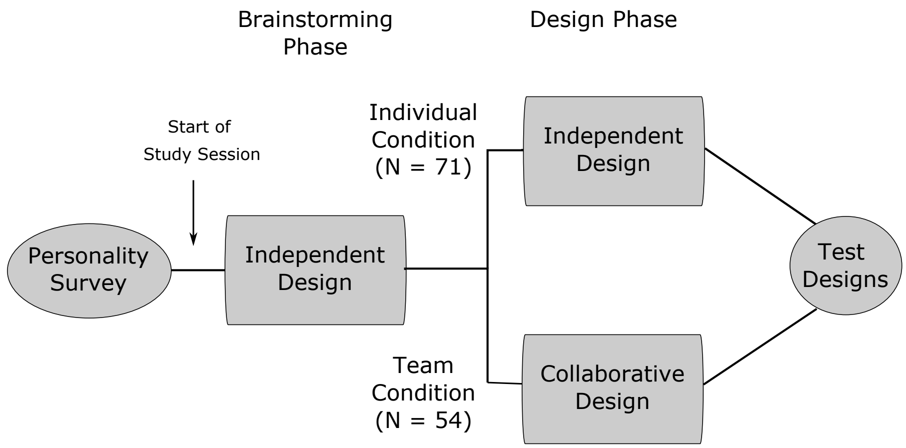
1.2 Our Approach

1.2.1 Rationale

The measure of influence we adopt in this study falls within the shift in opinion and behavior approach to measuring influence described above. Our study expands upon this approach by developing an objective measure of influence within the context of an open-ended design challenge. Our study leverages a virtual design game and is inspired by engineering design projects, allowing for a diversity of potential solutions that can vary in quality relative to a set of design criteria. This context, whereby the best solutions cannot be readily ascertained by participants, facilitated our measurement of how ideas influence decision-making (Jayles et al., 2017; Laughlin & Ellis, 1986). Our study design is particularly relevant to engineering design

Figure 1

Study Design

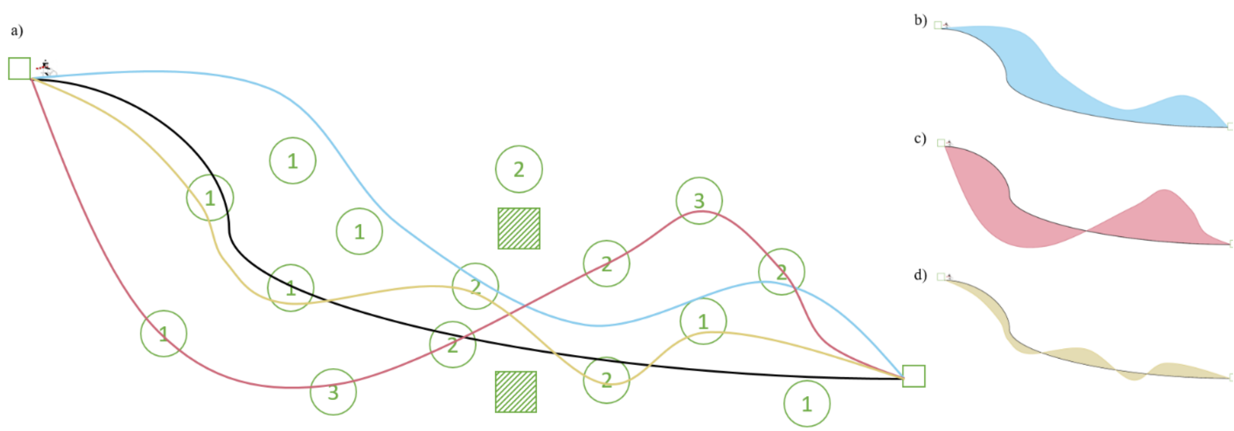


Note. The study design displays the Brainstorming and Design Phases across the Individual and Team Conditions. All participants completed a personality survey prior to the study and all designs were tested following the Design Phase. Teams comprised 162 individuals nested within 54 3-person teams.

teams, which emphasize teamwork, decision-making, prototyping, creativity, and performance (Han et al., 2022; Starkey et al., 2016; Zheng et al., 2018). Our study design is also relevant to temporary teams (i.e., groups of unfamiliar people with diverse skills who come together for a short duration to accomplish a complex task and then disband once the project is completed; Lv & Feng, 2021) and virtual teams, which have become more prevalent in the workplace over time (Bell & Kozlowski, 2002; Turesky et al., 2020). The task utilized the online program Line Rider (www.linerider.com), wherein a virtual sled rider rides down lines drawn by the user in a manner that approximates the rules of physics (See Methods). Participants were assigned the challenge of designing a simple track (i.e., consisting of a single continuous, non-overlapping line) that could enable the sled rider to achieve a goal defined by the experimenters. To better understand the effect of working on a team, we leveraged this platform to examine idea integration and performance across both teams (Team Condition) and individuals working independently (Individual Condition) (Han et al., 2022). Participants in both conditions experienced two distinct phases (Figure 1). In the Brainstorming Phase (Phase 1, 10 minutes), all participants designed a solution to the challenge independently. In the Design Phase (Phase 2, 15 minutes) participants produced a design that would be judged by the experimenters. Participants in the Team Condition worked in teams of three to come up with a single collaborative team design, and participants in the Individual Condition worked independently to produce their individual designs. Because tracks were composed of simple lines, the area between tracks represents a straightforward, quantitative measure of design similarity (Figure 2). We used this measure to infer the level of influence each participant's initial idea (i.e., the track they produced in the Brainstorming Phase) had over their subsequent design (i.e., the track they produced in the Design Phase), with smaller values indicating *more* influence (Bottger, 1984) (see Materials and Methods for more details). The advantages of our influence measure are that it is not impacted by team member perceptions or behaviors but rather reflects the quantifiable level of similarity between individuals' initial ideas and subsequent outputs. We did not manipulate participants' characteristics or behaviors as our study focuses on the natural associations between these variables and our influence measure (Bottger, 1984).

Figure 2

Design Challenge



Note. Figure 2a displays the background course (green) upon which participants designed their tracks. Stylized examples of team members' initial Brainstorming Phase tracks are displayed in blue, yellow, and red as well as a Design Phase (team) track in black. Figures 2b,c, and d show how the area between each Brainstorming Phase track and the Design Phase track can be calculated to obtain an objective measure of its influence over the team design (smaller areas mean greater influence)

1.2.2 Implementation

The current study provides an example of how our research approach can provide a precise and mechanistic understanding of the role influence plays in the emergence of collective intelligence within teams. Specifically, we demonstrate how our influence measure enables the calculation of the absolute (and relative) influence each team member's idea has over the team outcome which can serve as an objective basis for testing the predictors of influence and the effect influence processes have on team outcomes.

Performance and the Influence of Ideas in Teams and Individuals

Teams frequently display higher mean performance than individuals (Almaatouq et al., 2021; Goldman, 1965; Lorge et al., 1958). One explanation for this finding is that team members with greater expertise have an outsized influence over team outcomes (Bottger, 1984). For example, a previous study found that low-performing individuals paired with high-performers received the greatest boost in performance (compared to their initial individual-level performance), suggesting that the higher-performing partner had more influence over decision-making for the pair (Goldman, 1965). An alternative explanation for the role influence plays in enhanced team performance is that team members display more equal levels of influence, combining their ideas into new solutions that are more innovative or higher-performing than team members' initial ideas (Kohn et al., 2011; Paulus & Dzindolet, 2008). This is especially true for tasks where individuals possess distinct information that must be combined to come to an effective solution (Bedard et al., 1998; Toma & Butera, 2009). Indeed, some argue that how team members combine their ideas and other contributions is key to greater team performance and collective intelligence, even surpassing the role of team member intelligence (Engel et al., 2014; Haan et al., 2021).

Although the combined influence of team members' ideas can sometimes lead to greater team performance compared to that expected of individuals working independently (i.e., process gains), it can also lead to poorer team performance (i.e., process losses) (Bedard et al., 1998; Hackman & Morris, 1975). One way in which process losses may occur is when teams are influenced by ideas previously expressed by team members (Houlette et al., 2000) rather than continuing to explore the larger decision space (Bernstein et al., 2018). Consensus processes can also impede performance, whereby team members are more influenced by ideas better represented within the team, independent of their quality (De Dreu & West, 2001; Gigone & Hastie, 1997; Hackman & Morris, 1975). Furthermore, team members may gain influence over team outcomes through means other than expertise, such as dominance, which can have a negative impact on team expertise utilization and performance (Cheng et al., 2013; Sherf et al., 2018). Accordingly, the influence of others' ideas during teamwork has the potential to result in a lower likelihood of finding the optimal solution (Mason & Watts, 2012), and lower overall performance (Lorenz et al., 2011) compared to sets of individuals working independently.

Using our research approach, we first tested how the teams in our study performed compared to individuals— both those who worked independently throughout the study and the individual members of each team while they were working independently during the Brainstorming Phase. Comparing team with individual performance places our study within the context of other research on collective intelligence, and it is important for interpreting whether the mechanisms of idea integration we identified were associated with process gains or losses. We then conducted exploratory analyses to gain insight into the degree to which team member ideas influenced team outcomes in the Team Condition and how team members' levels of influence in the Team Condition compare to the influence individuals' ideas had over their final designs in the Individual Condition. By leveraging the quantitative nature of our data, we not only calculate individuals' levels of absolute and relative influence but also establish a theoretical baseline that enables us to test whether the level of individual influence observed was greater or lesser than that expected due to chance alone.

Team Member Characteristics, Influence, and Team Performance

It is through communication that team members combine their cognitive resources to reach a collective outcome (De Dreu et al., 2008). Speaking time is an interaction measure, which as described above, may be used as a proxy for influence. Speaking time is also frequently associated with positive perceptions of team member abilities and may thus correlate with measures of attributed influence (March, 1955). For example, speaking time is commonly associated with perceptions of expertise, competence, idea quality, influence, and leadership (Bales, 1953; Bottger, 1984; Li et al., 2019; Littlepage et al., 1995, 1995; MacLaren et al., 2020). Thus, several studies of collective intelligence have examined the role speaking time plays in its emergence. For example, when verbal (or work) contributions are centralized around a subset of team members, teams have been found to display lower quality outputs (Engel et al., 2014; Woolley et al., 2015), presumably because individuals who speak more prevent the utilization of other team members' expertise (Haan et al., 2021). There is evidence that the traits of central team members are also important considerations. For example, Sherf et al. (2018) found that when individuals with greater perceived voice (i.e., expression of work-related ideas) were higher in dominance, teams perceived lower team member expertise utilization and displayed lower performance, while such a relationship was not found when central team members were higher in reflectiveness (i.e., thoughtful, self-disciplined). As dominance has been tied to both perceived and objective measures of influence (Cheng et al., 2013), individuals who express relatively more work-related ideas and are also high in dominance may have additional influence over team outcomes.

In contrast to the perspective that team member communication quantity is central to the emergence of influence processes, other studies have emphasized the quality of information conveyed. For example, Bottger (1984) found that although speaking time was more highly correlated with perceived influence, expertise (high performance of initial individual ideas) was more highly correlated with objective measures of influence over team outcomes (which measured changes in team member responses between an initial independent phase and subsequent team phase). These findings are similar to those of March (1956), who found objective measures of influence to be independent of measures of interaction participation and attributed influence. Thus, some differences in findings across studies regarding the characteristics that are important for influence may be due to differences in the measure of influence used. However, there is also evidence that the quality of communicated information may interact with quantity to impact individual and team outcomes. For example, Ginter and Lindskold (1975) found that participants with high perceived expertise were more likely to be perceived as leaders despite their speaking time while those with low perceived expertise were more likely to be perceived as leaders when they spoke more. In addition, Bottger (1984) found that expertise is a greater predictor of objective influence when expertise and speaking time covary, and that teams display higher performance under these conditions compared to when expertise and speaking time are uncorrelated. Thus, individuals with relatively more expertise and speaking time may enjoy greater influence over team outcomes and also promote higher team performance.

The above studies suggest that the influence team members' ideas have over team decision-making can enhance or hinder collective intelligence according to a complex interplay between team members' ideas, behaviors, and traits (Becker et al., 2017). However, as these studies differ in whether or how they measure influence (i.e., self-report, interaction measurements, measurement of performance or behavior change), it remains difficult to assess the relationship between individual characteristics, the level of influence individuals have over team decision-making, and the impact these relationships have on team performance and collective intelligence. We use our research approach to conduct exploratory analyses that test the relationship between an individual's characteristics and the objective level of influence they have over team outcomes as measured by changes in team member designs between an initial independent phase and a subsequent team phase. The individual characteristics – or predictors – we considered were speaking time (an interaction measure that reflects the total time spent speaking during the Design Phase), dominance level (measured from a pre-study self-report survey), and expertise (the objectively measured performance of an individual's Brainstorming Phase track), as well as the interactions between 1) speaking time and dominance and 2) expertise and speaking time. We also controlled for an individual's work contributions (an interaction measure reflecting the total time spent controlling the shared screen during the Design Phase), given that participants could communicate and produce their designs by drawing on a shared screen. There is evidence that the distribution of work contributions across team members is related to the quality of team outcomes (Engel et al., 2014), making it important to control for this variable. In addition, we controlled for the objective similarity between team members' brainstorming ideas since brainstorming ideas that happen to be more similar to the ideas of other team members may have greater influence due to consensus processes (De Dreu & West, 2001; Gigone & Hastie, 1997; Hackman & Morris, 1975; March, 1956). Finally, we tested how the above characteristics relate to team performance by way of the team's most influential team member (Humphrey et al., 2009; Sherf et al., 2018). Based upon previous research (Bottger, 1984; Bunderson, 2003; Goldman, 1965), our expectation was that team performance would be associated with influential team members who have greater expertise, either through the effect of expertise alone or in combination with traits that heighten their influence.

1.3 Conclusion of Introduction

In summary, our study develops an objective measure of influence that facilitates understanding of the role influence processes play in team decision-making and the emergence of collective intelligence. Using this measure, our exploratory analyses demonstrate how ideas were combined within the teams in our study, what individual characteristics correlated with influence, and how these processes related to team performance, both across teams and relative to individuals. By quantitatively examining the objective influence individuals' ideas have over team decision-making, our approach can provide a more mechanistic understanding of the emergence of collective intelligence in teams (O'Bryan et al., 2020). In the Discussion, we consider how such approaches can be used to advance research within the field of collective intelligence.

MATERIALS AND METHODS

Participants

This research complied with the American Psychological Association Code of Ethics and was approved by the institution's Institutional Review Board. During the Summer of 2020 and Spring of 2021, the university where the study took place (located in the Southern United States) moved all research online due to the Covid-19 pandemic. Thus, this study took place within virtual meeting rooms (i.e., Zoom). Participants were compensated either with course credit or $10. We collected data from 84 individuals and 59 3-person virtual teams. We only extracted data from tracks that followed all study guidelines, and we only included data from individuals and teams for whom we had data for all tracks produced across both phases (N = 71 individuals and 162 individuals nested within N = 54 3-person teams).

The Line Rider Task

Following a training video, participants were given five minutes to engage in a structured training exercise during which they were familiarized with the Line Rider platform (www.linerider.com) and how the sled rider responds to various tracks. Training was followed by a 10-minute independent Brainstorming Phase and a 15-minute independent (Individual Condition) or collaborative (Team Condition) Design Phase (Figure 1). Participants began from a blank slate in the Design Phase and were not allowed to access their brainstorming tracks. Participants were also not allowed to test their designs themselves in the Brainstorming or Design Phases, although participants learned the performance of their Design Phase track upon completion of the Design Phase.

The design challenge involved designing a track on top of a background "course" (Figure 2). The sled rider could not interact with the background course but could interact with lines drawn by the participants. The challenge was to design a track that enabled the sled rider to reach a target at the bottom right-hand corner of the course as fast as possible while passing through as many checkpoints as possible and while avoiding crashing. Checkpoints were circles drawn on the course that had different point values written inside them. To gain the points assigned to a given checkpoint, the track had to pass through it. Tracks could not pass through the shaded squares. We designed the course to represent an open-ended task without one clear optimal solution. We incentivized participants by offering a $20 award per participant for the individual and team that produced the best-performing designs in the Design Phase across the Individual and Team Conditions, respectively. Tracks were collected from individuals and teams following the completion of each phase. All study sessions were video and audio recorded using the Zoom platform with automatic transcript generation enabled.

Track Data Processing

All tracks were visually assessed to determine whether they met the requirements laid out by the experimenters at the start of the study. The requirements included designing a track that was a continuous line, avoiding drawing a line where any section of the line was above or below another section of the line (e.g., no loops), and ensuring that the line did not pass through the shaded squares. All tracks were visually examined to determine whether there were any extraneous marks on the screen that were not part of the track that the rider interacted with. If present, these marks were manually removed. X-Y coordinates of all line segments comprising each track were imported into R (Version 2022.07.2) (R Core Team, 2022) using the jsonlite package (Ooms, 2014). We subsampled tracks at each integer x-coordinate, and we calculated the median y-coordinate for each x-coordinate to account for thick or irregular lines which may have more than one y-coordinate positioned close to one another.

Measures

Track Performance / Expertise

Our measure of track quality considers the number of points the track successfully passed through, the speed of the track, and the percentage of the course that was successfully completed. We counted points only for the checkpoints that the track passed through that were located to the left of either the finish line or the point where the rider crashed, whichever came first. If the rider crashed before reaching the end of the course, we marked the point where the rider crashed on the course and extracted the x-coordinates. We used this value to calculate the percentage of the course that the rider completed. Finally, we calculated the time it took the rider to reach the end of the course or crash, whichever came first. We compared this time to a default time, which we calculated as the time it took a rider to reach that same x-coordinate while riding down a straight diagonal line. Our measure of speed is the default time divided by the actual time so that faster tracks resulted in more points. We calculated the final score by multiplying the number of successful points, the speed ratio, and the percentage of the course completed. Brainstorming Phase track performance (which the experimenters calculated after the study was completed and which participants were unaware of) is also referred to as team member expertise.

Level of Influence

We calculated the level of influence that an individual's track in the Brainstorming Phase had over its associated track in the Design Phase by calculating the area between these two lines (rgeos package (Bivand & Rundel, 2023); Figure 2). We divided a given area by the maximum possible area between lines (i.e., forming a rectangle encompassing the entire design environment) resulting in a measure of the *normalized difference* between tracks. We interpret normalized difference values close to 0 to indicate Brainstorming Phase designs that had a high level of influence on the Design Phase track and normalized difference values closer to 1 to indicate lower levels of influence.

Speaking Time

We calculated the speaking time of participants in the Team Condition by extracting data from Zoom transcripts. Our Zoom account was set up to automatically record data to the cloud, including meeting audio and video (with timestamps) and an audio transcript. Transcripts record speaker identities, the start and end times of speaking bouts (herein referred to as speaking turns), and transcribed text. We derived our speaking time measure from transcripts by extracting speaker names and their speaking turns' start and end times using an R script (R Core Team, 2022). We divided each individual's speaking time by the duration of the team's interaction period in the Design Phase, which we calculated by subtracting the start time of the team's first speaking turn from the end time of their team's last speaking turn in the Design Phase. Due to data recording errors, we did not obtain automated speaking time data from three teams. Thus, analyses involving speaking time include data from $N = 51$ teams.

Screen Control Time

We calculated screen control time by coding the start and end time of the period in which each team member controlled the screen during the Design Phase and summing these durations for each team member (see Supplemental Materials). These measures were extracted by coding a Zoom-recorded video of the team's shared screen using the program ELAN (version 6.3). Assistants were trained to code both the first and last time a given team member moved the mouse on the shared screen after requesting access. If a participant controlled the screen during multiple bouts between which another team member controlled the screen, a start and end time was determined for each separate bout. The screen control time for a given team member represents the sum of the duration of all their screen control bouts. Like speaking time, we divided screen control time by the duration of the team's screen control period during the Design Phase, calculated as the end of the last screen control bout minus the beginning of the first screen control bout. Interrater reliability scores (ICC) for two pairs of coders who double-coded 17% of the data were .75 and .88, respectively.

Dominance

Before participating in the study, participants completed a questionnaire that assessed their personality and demographic characteristics. The current study only focuses on the trait of dominance, which was measured via the 11-item dominance scale from the International Personality Item Pool ($\alpha = .82$ (Goldberg, 1999)). Items were rated on a scale of 1 (very inaccurate) to 5 (very accurate). Example items include "Try to surpass others' accomplishments" and "Try to outdo others."

Brainstorming Idea Similarity

We calculated the mean normalized difference (See Level of Influence section above) between a given team member's Brainstorming Phase track and each of their two team members' Brainstorming Phase tracks. This measure reflects the mean similarity between a team member's design and those of their team members, with smaller values indicating greater similarity.

Statistical Analysis

Comparing Performance Across Teams and Individuals

All statistical analyses were executed in R. Due to the non-normality of the performance data, we used Wilcoxon Rank Sum tests to compare track performance between the Individual and Team Conditions in the Brainstorming Phase and the Design Phases. We used paired Wilcoxon Rank Sum tests to compare team performance in the Design Phase to their team's worst-, median-, and best-performing brainstorming track scores. Effect sizes were calculated by dividing the Z statistic by the square root of the sample size using the wilcox_effsize function within the rstatix package (Kassambara, 2023).

Distribution of Influence Within Teams

Due to the constraints of the experiment, it is possible that track designs could share similar properties, even if they were created entirely independently (i.e., without team member influence). Thus, we expected some similarity between tracks due to chance alone. To determine the distribution of normalized difference values that should be expected due to chance alone, we conducted a permutation analysis (Puga-Gonzalez et al., 2021) using the designs produced in our study. This enabled us to determine whether the level of influence participants' Brainstorming Phase designs had on their Design Phase designs differed significantly from chance. By comparing participants' Brainstorming and Design Phase tracks, we calculated the median observed levels of normalized difference and (for teams only) relative normalized difference — the normalized difference of each team member divided by the sum of the normalized differences in each team. We then compared these observed values to null distributions representing the absence of influence between Brainstorming and Design Phase tracks. We generated these distributions by permutating Design Phase tracks across teams (in the Team Condition) and individuals (in the Individual Condition) and recalculating the normalized differences (See Supplemental Materials for more information**).** Within teams, we then compared the median observed normalized difference and relative normalized difference values displayed by the most, intermediate, and least influential ideas of each team to the distribution of median values expected due to chance alone within each influence category. For individuals, we compared median observed levels of normalized difference across all individuals to those expected due to chance.

In addition, we used Wilcoxon Rank Sum tests to examine how the normalized difference values observed between team member Brainstorming Phase tracks and Design Phase tracks in the Team Phase (divided into influence categories) compare to the normalized difference values observed in the Individual Condition. Effect sizes were calculated as described above. These analyses enabled us to determine how the influence of team members compared to that of individuals working independently.

Characteristics Associated with Influence

We used generalized linear mixed-effects models to examine the characteristics associated with the level of influence a team member's Brainstorming Phase track had over the team's Design Phase track. The predictor variables we considered were an individual's expertise, speaking time, dominance level, work contributions, and idea similarity. We also assessed the interactions between 1) speaking time and dominance and 2) expertise and speaking time. Due to differences in scale across model variables, all variables were normalized by subtracting the sample mean and dividing by the standard deviation using the scale function in R (R Core Team, 2022). Although all predictor variables in this analysis were measured at the individual level (Level 1), these individuals were grouped into teams. Thus, we use multilevel modeling with a random effect for Team to account for variation of the intercept between teams (i.e., "random intercept model"). Since the response variable was between 0 and 1, we used the glmmTMB function in R (Brooks et al., 2017) with a beta family and logit link function to fit the model. We verified model fit using the DHARMa package (Hartig, 2022).

Variation in Performance Across Teams

We tested how the characteristics of central team members (the most influential team member of each team) impacted the performance of teams' Design Phase tracks. Team tracks tended to closely match their most influential track, with a median [IQR] normalized difference value of .049 [.026 - .092]. However, the distribution of normalized difference values for these tracks ranged from .0091 – .32 (Supplemental Figure 1a) with 0 indicating a perfect match and 1 indicating the maximum possible difference between tracks in the course. Because small differences in track design could result in big differences in performance, particularly if a change resulted in the rider crashing before reaching the end of the course, a high-quality influential brainstorming track is not guaranteed to lead to a high-quality team design. Since track performance in the design phase was zero-inflated (e.g., due to some tracks that failed before the rider passed through any checkpoints), we rounded performance scores to the nearest whole number and used a zero-inflated linear model to model our performance data. We then focused on all (scaled) characteristics of a team's most influential team member (described above) to determine how central team members' characteristics related to variation in performance across teams. In addition to analyzing variation in overall performance, we examined variation in team performance relative to each team's median-performing brainstorming score. For this analysis, the response variable was team performance in the Design Phase minus the performance of the team's median-performing Brainstorming Phase track. This measure was normally distributed, so we used linear regression for this analysis. In addition, we group-mean-centered the characteristics of teams' most influential team members to reflect the value of the central individuals' characteristics (e.g., expertise, speaking time, dominance) relative to their team members.

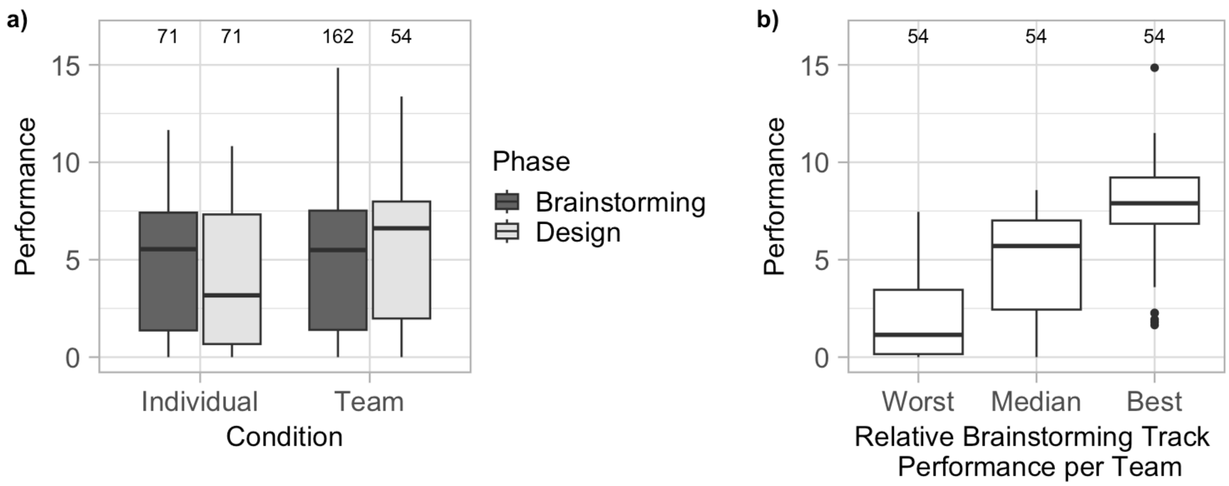
RESULTS

Comparing Performance Across Individuals and Teams

There was no significant difference between the performance of Team and Individual Condition tracks in the Brainstorming Phase, when all participants worked independently ($r = 0.016$, $W = 5867$, $p = .81$, Figure 3a). However, Design Phase tracks performed significantly better in the Team Condition, when team members worked together, compared to the Individual Condition when individuals worked independently ($r = .21$, $W = 2378.5$, $p = .022$, Figure 3a). Figure 3b displays the performance of teams' worst, median, and best-performing Brainstorming Phase tracks. Comparing these values to the performance of Design Phase tracks in the Team Condition, we found that teams' Design Phase tracks performed significantly better than their

Figure 3

Comparison of Performance



Note. a) Comparison of track performance across individuals and teams during the Brainstorming and Design Phases. Participants worked independently in the Brainstorming Phase across the Team and Individual Conditions. Participants worked together in teams of 3 in the Team Condition during the Design Phase. b) Breakdown of the worst-, median- and best-performing Brainstorming Phase tracks per team. Values at the top of the graphs represent sample sizes.

worst-performing Brainstorming Phase track (r = .68, $W = 1279$, $p < .001$), no differently than their median-performing Brainstorming Phase track (r = .21, $W = 918$, $p = 0.13$), and significantly worse than their best-performing Brainstorming Phase track (r = .45, $W = 355$, $p < .001$; Figure 3b). Thus, teams performed better than individuals when team members worked together but did not match the performance of their best idea.

Distribution of Influence within Teams

The median [IQR] levels of normalized difference between the most, intermediate, and least influential team member brainstorming tracks and their team's track were .049 [.026 - .092], .14 [.089 - .25], and .29 [.16 - .36], respectively, with 0 indicating a perfect match and 1 indicating the maximum possible difference between tracks (Supplemental Figure 1). When these median values were compared to those expected due to chance alone, we found that the most influential team member of each team ($p < .001$) and the team member of intermediate influence ($p < .001$) had lower normalized difference than expected, and thus greater influence (Supplemental Figure 3a). Observed values of normalized difference for the least influential team member of each team did not significantly differ from chance ($p = .086$). When examining team members' relative normalized difference, the median [IQR] values for the most, intermediate, and least influential

Table 1

Variable	M	SD	1	2	3	4	5
1. Normalized Difference	0.17	0.13					
2. Expertise	4.70	3.36	-.10 [-.25, .06]				
3. Idea Similarity	0.21	0.09	.38** [.24, .51]	.13 [-.03, .28]			
4. Speaking Time	0.15	0.08	-.08 [-.24, .08]	.05 [-.11, .21]	.00 [-.16, .16]		
5. Screen Control Time	0.27	0.30	-.26** [-.40, -.11]	.05 [-.11, .21]	-.04 [-.20, .12]	.37** [.23, .50]	
6. Dominance	2.71	0.54	-.07 [-.22, .09]	.05 [-.11, .21]	.08 [-.08, .23]	.06 [-.10, .22]	.19* [.04, .34]

Note. N = 153 team members within 51 teams. M and SD represent mean and standard deviation, respectively. Values in square brackets indicate the 95% confidence interval for each correlation. Influence and idea similarity are reverse coded, with smaller values indicating higher levels of influence and similarity, respectively. * $p < .05$. ** $p < .01$.

team members were 0.14 [.060 - .19], 0.33 [.27 - .39], and 0.52 [.46 - .61], where a value of 0.33 for all team members would represent equal influence (Supplemental Figure 2). When these median values were compared to those expected due to chance alone, we found that the most influential team member of each team had lower relative normalized difference than expected ($p < .001$) (and thus greater relative influence) and the least influential member of each team had greater relative normalized difference than expected ($p < .001$) (and thus less relative influence; Supplemental Figure 3c). Observed values of relative normalized difference for team members of intermediate influence did not differ significantly from chance ($p = .38$).

When examining the Individual Condition, we found that individuals had median [IQR] levels of normalized influence of .069 [.030 - .18] between their Brainstorming Phase and Design Phase tracks (Supplemental Figure 1d). These values were significantly lower than expected ($p < .001$; Supplemental Figure 3b) indicating that individuals' Brainstorming Phase tracks had greater influence on Design Phase tracks than expected due to chance. On average, an individual's level of influence was not significantly different from a team's most influential team member (r = .17, W = 1529, p = .053) but was significantly greater than that displayed by team members with intermediate (r = .31, W = 2603, $p < .001$) and lowest (r = .55, W = 3153, $p < .001$) levels of influence (Supplemental Figure 1).

Characteristics Associated with Influence

Means, standard deviations, and correlations among individual characteristics are displayed in Table 1. As none of the interaction terms in our model of individual influence were significant, we removed them from the final model. The results of this reduced model are reported below, and all regression model results can be found in the Individual Influence column in Table

Table 2

	Individual Influence[a]		Team Performance[b]		Relative Team Performance[c]	
Variables	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
Intercept	-1.65 (.068)***	-1.65 (.068)***	1.80 (.077)***	1.79 (.074)***	.60 (.74)	.69 (.70)
Expertise	-.16 (.065)*	-.16 (.065)*	.18 (.062)**	.19 (.058)**	.37 (.86)	.45 (.83)
Speaking time	.046 (.069)	.044 (.066)	-.13 (.067)*	-.11 (.060)†	-.63 (.83)	-.50 (.75)
Dominance	-.0077 (.066)	-.0085 (.066)	-.042 (.077)	-.047 (.072)	-.93 (.69)	-.89 (.68)
Expertise x Speaking time	-.015 (.057)		.040 (.042)		.048 (.82)	
Speaking time x dominance	.0052 (.059)		-.044 (.085)		.59 (.82)	
Work contributions	-.26 (.072)***	-.26 (.072)***	.15 (.071)*	.16 (.066)*	1.01 (.69)	.98 (.68)
Idea similarity	.21 (.058)***	.21 (.058)***	-.044 (.053)	-.045 (.053)	1.04 (1.25)	1.14 (1.16)

Note. Model coefficients (with SE in parentheses) for models of individual influence, team performance, and team performance relative to the performance of the team's median Brainstorming Phase track score. a) N = 153 individuals nested within 51 groups. All variables were scaled before being entered into the model. We included Team as a random effect due to the nested nature of the data. b) N = 51 teams. Variables reflect the characteristics of the team's most influential team member. All variables were scaled before being entered into the model. c) N = 51 teams. Variables reflect the characteristics of the team's most influential team member. All variables were scaled and group mean centered before being entered into the model. † $p < .10$; * $p < .05$; ** $p < 0.01$; *** $p < .001$

2. The normalized difference between a team member's Brainstorming Phase track and their team's Design Phase track was significantly negatively related to their expertise (B = -.16, SE = .065, p = .011) and work contributions (B = -.26, SE = .072, p < .001). These results indicate that those with greater expertise and work contributions had greater influence over the team solution. A team member's idea similarity was positively associated with their idea's normalized difference from the team track (B = .21, SE = .058, p < .001) meaning that individuals who had more similar designs to those of their team members had more influence over the team design. A team member's speaking time (B = .044, SE = .066, p = .50) and dominance level (B = -.0085, SE = .066, p = .90) did not have a significant relationship with their brainstorming design's normalized difference from the team track. However, as seen in the correlation table (Table 1), there were positive correlations between screen control time and both dominance and speaking time.

Central Individuals and Performance

As none of the interaction terms in our models of team performance were significant, we removed them from the final model. The results of these reduced models are reported below and all regression model results can be found in the Team Performance and Relative Team Performance columns in Table 2. We found that the central team member's expertise (B = 0.19, SE = 0.058, p = .0012) and their work contributions (B = 0.16, SE = 0.066, p = .013) were positively correlated with team performance, with no effect of their speaking time (B = -0.11, SE = 0.060, p = 0.071), dominance level (B = -0.047, SE = 0.072, p = 0.51), or the similarity of their idea to others' (B = -0.045, SE = 0.053, p = 0.39). Thus, teams performed best when influential team members had high-quality ideas and when they had greater control over work contributions. None of the group-mean-centered predictors were significantly correlated with relative team performance (i.e., how a team performed relative to their median-performing brainstorming track score; Table 2).

DISCUSSION

This study's unique contribution is the expansion of methods that can be used to examine mechanisms by which individuals' ideas influence team outcomes. To this end, we developed an objective measure of influence that focuses on the measurement of shifts in designs over time, and we demonstrate how this approach can be used to test hypotheses related to team member influence and team performance. Expanding upon previous studies that have focused on objective measures of influence (Bottger, 1984; Littlepage et al., 1995; March, 1956), we examined performance and influence across both teams and individuals working independently and developed methods for examining how multiple ideas were integrated into a single team solution. Our exploratory results demonstrated that 1) some team members exhibited greater influence than expected, at the expense of others' influence, 2) team members with greater expertise, who made more work contributions, and whose ideas were more similar to other team members' ideas had greater influence over team outcomes, and 3) teams performed best when the most influential team member had a high-quality brainstorming idea and made more work contributions. We interpret these findings within the context of teams that performed better than independent individuals but no better than their median team member and worse than their best team member. Overall, we demonstrate how using objective measures of influence can aid in clarifying the mechanisms by which collective intelligence can emerge and discuss how these approaches can be leveraged in future studies of group decision-making and collective intelligence.

In line with previous studies (Lorge et al., 1958; Salas et al., 2018), the teams in our study performed better, on average, than individuals working independently even though there was no difference in brainstorming track performance between participants in the Team and Individual Conditions. Thus, the teams in our study displayed a collective advantage compared to individuals who continued to work alone. However, since teams did not perform any differently than the median performance of their team members' brainstorming ideas and underperformed compared to their highest-performing idea, teams did not display process gains that enabled them to perform better than the individual members of their team. Although these findings contradict some reports of teams performing better than their average team member (Lorge et al., 1958) and even their best team member (Michaelsen & Watson, 1989; Nemiroff & King, 1975), they are in line with other studies that have found that teams often fall short of these markers of a collective advantage (Hackman, 2002; O'Neill & Salas, 2018). Comparing team to individual performance (both individuals working independently and individuals within a given team) is an important step for identifying teams that display process gains and those that do not (Han et al., 2022). In addition, this approach can be used to identify tasks that promote the emergence of process gains or tasks that require them for teams to perform well. For example, it is well-known that hidden-profile tasks require team members to integrate distinct pieces of information to produce high-quality outcomes (Toma & Butera, 2009). By building these characteristics into a task designed around the Line Rider platform, researchers can test when and under what conditions teams display process gains (or losses).

A key contribution of our study is that we combine our assessment of a team's collective advantage with analyses of how ideas were integrated within these teams. We found that teams in our study gravitated towards the ideas of some team members and away from others. The fact that team members did not merge their ideas more equally could be because some team members' ideas could be incompatible with one another, such as a track following a route towards the top of the course and another towards the bottom. This is because the simulated rules of physics within the game limited how the rider moved along tracks. Indeed, studies of engineering design teams have found that teams tend to select feasible designs over original designs to reduce uncertainty (Rietzschel et al., 2010; Starkey et al., 2016). Thus, teams in our study may have avoided more creative ideas in favor of ones they anticipated would perform effectively. By building upon the approach used in our study to test different methods of idea integration, future studies could test both the task characteristics (i.e., how well full or partial ideas can be merged) and team interaction processes that promote more or less equal methods of idea integration. For example, Woolley et al. (2015) found that teams that display more equal speaking turns across team members perform better, presumably because they more fully take into account team member expertise. Using our approach for analyzing idea integration across team members, it could be possible to test whether teams that display more equal speaking turn patterns tend to integrate their ideas more equally and how this relationship depends upon both task characteristics and the distribution of expertise within the team.

The level of influence the most influential team member displayed was not significantly different from the influence an individual's Brainstorming Phase track had on their Design Phase track when working independently. Thus, even though teams were strongly influenced by some team members' initial ideas, teams did not appear to explore the broader decision space any less than individuals working independently. These findings contradict previous studies that have found that the influence others' ideas have on individual and collective decision-making can impede exploration of the larger decision space (Bernstein et al., 2018; Lorenz et al., 2011). Nevertheless, since the teams in our study had three team member designs to choose from, they likely benefitted by having access to a greater diversity of design options to select from compared to individuals who were strongly influenced by only their own design. By using our approach to compare the level of influence displayed by the most influential team member to the level of influence displayed by individuals working independently, future studies can test the conditions under which teams explore the decision space less (or more) thoroughly than individuals working independently. For example, studies could explore the effect of using different concept selection tools (Zheng et al., 2018) or the emphasis on convergent versus divergent thinking (Hirshfield & Koretsky, 2021) on teams' tendencies to explore the decision space.

In line with previous studies using objective measures of influence as an outcome (Bottger, 1984), we found that the quality of a team member's idea was positively correlated with its objective level of influence over the team outcome. However, as influence within teams was also correlated with team member work contributions and idea similarity, the teams in our study did not appear to base their decision-making purely upon team member expertise alone. Although greater work contributions by influential team members were positively correlated with team performance, a team member's work contributions were not correlated with their expertise. Rather, work contributions were correlated with dominance and speaking time even though these variables were not directly associated with influence themselves. Since team members could communicate their ideas both verbally and by drawing on the shared screen, these dual modalities may have reduced the effect of speaking time on influence. On the other hand, some studies have found that even though speaking time is positively associated with *perceived* influence and leadership, it does not correspond with *actual* influence (Bottger, 1984; MacLaren et al., 2020). Thus, our study's focus on objective measures of influence may be another reason why we did not find an effect of speaking time. Focusing on objective measures of influence, such as that developed in our study, can therefore help to clarify the ways in which team member contributions are integrated into team outcomes and enable the differentiation between perceived and actual influence. In future studies, our research approach can be used to test the conditions under which given individual characteristics (e.g., expertise, speaking time) correlate with influence, such as by manipulating the importance of expertise to a given task and the number of modalities that are available for communication. For example, the communication modality in the Line Rider task could be restricted to only verbal communication by having team members discuss their ideas verbally and give verbal instructions to a neutral participant controlling the screen.

One limitation of our approach is that our measure of team member expertise reflects the performance of their design in the Brainstorming Phase (which the participants could not test during the study) and did not take into account the track's future potential (Girotra et al., 2010). Furthermore, since small errors in the track (e.g., bumps, gaps) had the potential to impede performance, our measure may have underestimated the quality of some tracks. Another limitation of our approach is that our measure of influence only takes into account an individual's initial idea. Although individuals may come up with additional ideas as they work with their team members, our approach currently does not take these additional inputs into account. However, as team members in our study tended to strongly match the initial design of one team member, capturing additional ideas may not be as important as focusing on team members' initial ideas. Another point to consider is that the creator of a given idea may not necessarily be the one who recognizes its value and promotes it within the team (Toh & Miller, 2015). In order to investigate the importance of this facet, content analysis could be considered to determine which team members support which ideas during verbal interactions. In this way, combining our objective measure of influence with additional influence measures (i.e., measures of interaction, attributed influence) may help to provide a more complete picture of influence processes.

Our study population consisted of students working in a controlled environment. Since the participants were being actively monitored and were provided specific instructions, this context likely resulted in strong situation norms that may not exist in the real world (Mischel, 1973). Additionally, this study utilized ad-hoc teams, defined as teams that consist of members with no prior working experience (Altschuller & Benbunan-Fich, 2010) which are assembled for a short time period after which they disband (Feitosa et al., 2020). In such teams, there are generally fewer opportunities to develop a team identity, shared mental models, and trust (White et al., 2018). Additional work is needed to expand understanding of team processes in teams that work together for longer periods (Saunders & Ahuja, 2006). Thus, future studies can expand upon the approaches described here by adapting these methods to teams interacting over longer periods of time and in settings that are more relevant to real-world employees. In addition, future extensions can take into account additional stages of the engineering design process, such as the testing and further refinement of prototypes.

By developing methods for measuring and analyzing objective measures of influence within the context of an open-ended design task, our study provides novel insight into how team members' ideas, behaviors, and traits impact their influence over team decision-making. By improving our ability to measure how team member ideas influence team decision-making we can not only better understand the team processes associated with the emergence of collective intelligence but also gain a more precise understanding of how to improve these processes. In addition, by moving away from more subjective measures of influence, such as self-report, and towards more objective measures, we may enhance our ability to compare the influence processes underlying collective intelligence and other forms of enhanced group-level abilities observed across biological, social, and artificial systems (O'Bryan et al., 2020).

ACKNOWLEDGEMENTS

We would like to thank the undergraduate research assistants in the Adult Skills and Knowledge Lab who aided the development of the study task, the facilitation of experiments, and the coding of experimental data. Funding for this project was provided by Microsoft Research (Microsoft Productivity Research Grant) and the National Science Foundation (Award Number: 1910117).

REFERENCES

Almaatouq, A., Alsobay, M., Yin, M., & Watts, D. J. (2021). Task complexity moderates group synergy. *Proceedings of the National Academy of Sciences*, *118*(36), e2101062118. https://doi.org/10.1073/pnas.2101062118

Altschuller, S., & Benbunan-Fich, R. (2010). Trust, Performance, and the Communication Process in Ad Hoc Decision-Making Virtual Teams. *Journal of Computer-Mediated Communication*, *16*(1), 27–47. https://doi.org/10.1111/j.1083-6101.2010.01529.x

Bales, R. F. (1953). The equilibrium problem in small groups. *Working Papers in the Theory of Action*, *111*(161), 3–45.

Becker, J., Brackbill, D., & Centola, D. (2017). Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, *114*(26), E5070–E5076. https://doi.org/10.1073/pnas.1615978114

Bedard, J. C., Biggs, S. F., Maroney, J. J., & Johnson, E. N. (1998). Sources of Process Gain and Loss From Group Interaction in Performance of Analytical Procedures. *Behavioral Research in Accounting*, *10*, 207–239.

Bell, B. S., & Kozlowski, S. W. J. (2002). A typology of virtual teams: Implications for effective leadership. *Group & Organization Management*, *27*(1), 14–49.

Bernstein, E., Shore, J., & Lazer, D. (2018). How intermittent breaks in interaction improve collective intelligence. *Proceedings of the National Academy of Sciences*, *115*(35), 8734–8739. https://doi.org/10.1073/pnas.1802407115

Bivand, R., & Rundel, C. (2023). *Rgeos: Interface to Geometry Engine—Open Source ('GEOS')* [Computer software]. https://r-forge.r-project.org/projects/rgeos/ https://libgeos.org http://rgeos.r-forge.r-project.org/index.html

Bonabeau, E., Dorigo, M., & Theraulaz, G. (1999). *Swarm Intelligence: From Natural to Artificial Systems*. Oxford University Press.

Bottger, P. C. (1984). Expertise and air time as bases of actual and perceived influence in problem-solving groups. *Journal of Applied Psychology*, *69*(2), 214–221. https://doi.org/10.1037/0021-9010.69.2.214

Brooks, M. E., Kristensen, K., Benthem, K. J. van, Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378–400.

Bunderson, J. S. (2003). Recognizing and Utilizing Expertise in Work Groups: A Status Characteristics Perspective. *Administrative Science Quarterly*, *48*(4), 557–591. https://doi.org/10.2307/3556637

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology*, *104*(1), 103–125. https://doi.org/10.1037/a0030398

De Dreu, C. K. W., Nijstad, B. A., & van Knippenberg, D. (2008). Motivated Information Processing in Group Judgment and Decision Making. *Personality and Social Psychology Review*, *12*(1), 22–49. https://doi.org/10.1177/1088868307304092

De Dreu, C. K. W., & West, M. A. (2001). Minority dissent and team innovation: The importance of participation in decision making. *Journal of Applied Psychology*, *86*(6), 1191–1201. https://doi.org/10.1037/0021-9010.86.6.1191

Deuling, J. K., Denissen, J. J. A., Van Zalk, M., Meeus, W., & Aken, M. V. (2011). Perceived influence in groups over time: How associations with personality and cognitive ability can change over time. *Journal of Research in Personality*, *45*(6), 576–585. https://doi.org/10.1016/j.jrp.2011.07.005

Engel, D., Woolley, A. W., Jing, L. X., Chabris, C. F., & Malone, T. W. (2014). Reading the Mind in the Eyes or Reading between the Lines? Theory of Mind Predicts Collective Intelligence Equally Well Online and Face-To-Face. *PLoS ONE*, *9*(12), e115212–e115212. https://doi.org/10.1371/journal.pone.0115212

Feitosa, J., Grossman, R., Kramer, W. S., & Salas, E. (2020). Measuring team trust: A critical and meta-analytical review. *Journal of Organizational Behavior*, *41*(5), 479–501. https://doi.org/10.1002/job.2436

Gigone, D., & Hastie, R. (1997). The impact of information on small group choice. *Journal of Personality and Social Psychology*, *72*(1), 132–140. https://doi.org/10.1037/0022-3514.72.1.132

Gintner, G., & Lindskold, S. (1975). Rate of Participation and Expertise as Factors Influencing Leader Choice. *Journal of Personality and Social Psychology*, *32*(6), 1085–1089. https://doi.org/10.1037/0022-3514.32.6.1085

Girotra, K., Terwiesch, C., & Ulrich, K. T. (2010). Idea Generation and the Quality of the Best Idea. *Management Science*, *56*(4), 591–605. https://doi.org/10.1287/mnsc.1090.1144

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe*, *7*(1), 7–28.

Goldman, M. (1965). A comparison of individual and group performance for varying combinations of initial ability. *Journal of Personality and Social Psychology*, *1*(3), 210–216. https://doi.org/10.1037/h0021874

Haan, K.-W., Riedl, C., & Woolley, A. (2021). Discovering Where We Excel: How Inclusive Turn-Taking in Conversation Improves Team Performance. *Companion Publication of the 2021 International Conference on Multimodal Interaction*, 278–283. https://doi.org/10.1145/3461615.3485417

Hackman, J. R. (2002). Why Teams Don't Work. In R. S. Tindale, L. Heath, J. Edwards, E. J. Posavac, F. B. Bryant, Y. Suarez-Balcazar, E. Henderson-King, & J. Myers (Eds.), *Theory* *and Research on Small Groups* (pp. 245–267). Springer. https://doi.org/10.1007/0-306-47144-2_12

Hackman, J. R., & Morris, C. G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. *Advances in Experimental Social Psychology*, *8*, 45–99. https://doi.org/10.1016/S0065-2601(08)60248-8

Han, J., Park, D., Hua, M., & Childs, P. R. N. (2022). Is group work beneficial for producing creative designs in STEM design education? *International Journal of Technology and Design Education*, *32*(5), 2801–2826. https://doi.org/10.1007/s10798-021-09709-y

Hartig, F. (2022). *DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models*. http://florianhartig.github.io/DHARMa/

Hirshfield, L. J., & Koretsky, M. D. (2021). Cultivating creative thinking in engineering student teams: Can a computer-mediated virtual laboratory help? *Journal of Computer Assisted Learning*, *37*(2), 587–601. https://doi.org/10.1111/jcal.12509

Houlette, M., Muzzy, E. L., & Sawyer, J. E. (2000). Information sharing and integration in culturally diverse cross-functional groups. *JE Sawyer (Chair), Decision-Group Composition, Structure and Processes: Applications of the Hidden Profile Paradigm to Group and Organizational Research. Symposium Presented at the Annual Meeting of the Academy of Management, OB and MOC Divisions, Toronto, Canada*.

Humphrey, S. E., Morgeson, F. P., & Mannor, M. J. (2009). Developing a theory of the strategic core of teams: A role composition model of team performance. *Journal of Applied Psychology*, *94*(1), 48–61. https://doi.org/10.1037/a0012997

Jayles, B., Kim, H., Escobedo, R., Cezera, S., Blanchet, A., Kameda, T., Sire, C., & Theraulaz, G. (2017). How social information can improve estimation accuracy in human groups. *Proceedings of the National Academy of Sciences*, *114*(47), 12620–12625. https://doi.org/10.1073/pnas.1703695114

Kassambara, A. (2023). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* [Computer software]. https://CRAN.R-project.org/package=rstatix

Kihlstrom, J. F., Eich, E., Sandbrand, D., & Tobias, B. A. (1999). Emotion and Memory: Implications for Self-Report. In A. A. Stone, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The Science of Self-report. Implications for Research and Practice* (1st ed., pp. 81–100). Erlbaum.

Kohn, N. W., Paulus, P. B., & Choi, Y. (2011). Building on the ideas of others: An examination of the idea combination process. *Journal of Experimental Social Psychology*, *47*(3), 554–561. https://doi.org/10.1016/j.jesp.2011.01.004

Kozlowski, S. W. J., & Chao, G. T. (2018). Unpacking team process dynamics and emergent phenomena: Challenges, conceptual advances, and innovative methods. *American Psychologist*, *73*(4), 576–592. https://doi.org/10.1037/amp0000245

Laughlin, P. R., & Ellis, A. L. (1986). Demonstrability and Social Combination Processes on Mathematical Intellective Tasks. *Journal of Experimental and Social Psychology*, *22*(3), 177–189. https://doi.org/10.1016/0022-1031(86)90022-3

Li, H., Yuan, Y. C., Bazarova, N. N., & Bell, B. S. (2019). Talk and Let Talk: The Effects of Language Proficiency on Speaking Up and Competence Perceptions in Multinational Teams. *Group and Organization Management*, *44*(5), 953–989. https://doi.org/10.1177/1059601118756734

Littlepage, G. E., Schmidt, G. W., Whisler, E. W., & Frost, A. G. (1995). An Input-Process-Output Analysis of Influence and Performance in Problem-Solving Groups. *Journal of Personality and Social Psychology*, *69*(5), 877–889. https://doi.org/10.1037/0022-3514.69.5.877

Lorenz, J., Rauhut, H., Schweitzer, F., & Helbing, D. (2011). How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences*, *108*(22), 9020–9025. https://doi.org/10.1073/pnas.1008636108

Lorge, I., Fox, D., Davitz, J., & Brenner, M. (1958). A survey of studies contrasting the quality of group performance and individual performance, 1920-1957. *Psychological Bulletin*, *55*(6), 337–372. https://doi.org/10.1037/h0042344

Lv, M., & Feng, S. (2021). Temporary teams: Current research focus and future directions. *Quality & Quantity*, *55*(1), 1–18. https://doi.org/10.1007/s11135-020-00990-y

MacLaren, N. G., Yammarino, F. J., Dionne, S. D., Sayama, H., Mumford, M. D., Connelly, S., Martin, R. W., Mulhearn, T. J., Todd, E. M., Kulkarni, A., Cao, Y., & Ruark, G. A. (2020). Testing the babble hypothesis: Speaking time predicts leader emergence in small groups. *Leadership Quarterly*, *31*(5), 101409. https://doi.org/10.1016/j.leaqua.2020.101409

March, J. G. (1955). An Introduction to the Theory and Measurement of Influence. *American Political Science Review*, *49*(2), 431–451. https://doi.org/10.2307/1951813

March, J. G. (1956). Influence Measurement in Experimental and Semi-Experimental Groups. *Sociometry*, *19*(4), 260–271. https://doi.org/10.2307/2785767

Mason, W., & Watts, D. J. (2012). Collaborative learning in networks. *Proceedings of the National Academy of Sciences*, *109*(3), 764–769. https://doi.org/10.1073/pnas.1110069108

Mast, M. S. (2002). Dominance as Expressed and Inferred Through Speaking Time. A Meta-Analysis. *Human Communication Research*, *28*(3), 420–450. https://doi.org/10.1111/j.1468-2958.2002.tb00814.x

Michaelsen, L. K., & Watson, W. E. (1989). A Realistic Test of Individual Versus Group Consensus Decision Making. *Journal of Applied Psychology*, *74*(5), 834–839. https://doi.org/10.1037/0021-9010.74.5.834

Mischel, W. (1973). Toward a cognitive social learning reconceptualization of personality. *Psychological Review*, *80*(4), 252–283.

Nemiroff, P. M., & King, D. C. (1975). Group Decision-Making Performance as Influenced by Consensus and Self-Orientation. *Human Relations*, *28*(1), 1–21. https://doi.org/10.1177/001872677502800101

O'Bryan, L., Beier, M., & Salas, E. (2020). How approaches to animal swarm intelligence can improve the study of collective intelligence in human teams. *Journal of Intelligence*, *8*(1), 9. https://doi.org/10.3390/jintelligence8010009

O'Neill, T. A., & Salas, E. (2018). Creating high performance teamwork in organizations. *Human Resource Management Review*, *28*(4), 325–331. https://doi.org/10.1016/j.hrmr.2017.09.001

Ooms, J. (2014). The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. *arXiv:1403.2805 [Stat.CO]*. https://arxiv.org/abs/1403.2805

Paulus, P. B., & Dzindolet, M. (2008). Social influence, creativity and innovation. *Social Influence*, *3*(4), 228–247. https://doi.org/10.1080/15534510802341082

Puga-Gonzalez, I., Sueur, C., & Sosa, S. (2021). Null models for animal social network analysis and data collected via focal sampling: Pre-network or node network permutation? *Methods in Ecology and Evolution*, *12*(1), 22–32. https://doi.org/10.1111/2041-210X.13400

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. https://www.R-project.org/

Riecken, H. W. (1958). The Effect of Talkativeness on Ability to Influence Group Solutions of Problems. *Sociometry*, *21*(4), 309–321. https://doi.org/10.2307/2785795

Riedl, C., Kim, Y. J., Gupta, P., Malone, T. W., & Wooley, A. W. (2021). Quantifying collective intelligence in human groups. *Proceedings of the National Academy of Sciences*, *118*(21), 1–5. https://doi.org/doi:10.1073/pnas.2005737118

Rietzschel, E. F., Nijstad, B. A., & Stroebe, W. (2010). The selection of creative ideas after individual idea generation: Choosing between creativity and impact. *British Journal of Psychology*, *101*(1), 47–68. https://doi.org/10.1348/000712609X414204

Salas, E., Reyes, D. L., & McDaniel, S. H. (2018). The science of teamwork: Progress, reflections, and the road ahead. *American Psychologist*, *73*(4), 593–600. https://doi.org/10.1037/amp0000334

Saunders, C. S., & Ahuja, M. K. (2006). Are All Distributed Teams the Same? Differentiating Between Temporary and Ongoing Distributed Teams. *Small Group Research*, *37*(6), 662–700. https://doi.org/10.1177/1046496406294323

Seeley, T. D., & Buhrman, S. C. (1999). Group decision making in swarms of honey bees. *Behavioral Ecology and Sociobiology*, *45*, 19–31. https://doi.org/10.1007/s002650050536

Sherf, E. N., Sinha, R., Tangirala, S., & Awasty, N. (2018). Centralization of member voice in teams: Its effects on expertise utilization and team performance. *Journal of Applied Psychology*, *103*(8), 813–827. https://doi.org/10.1037/apl0000305

Starkey, E., Toh, C. A., & Miller, S. R. (2016). Abandoning creativity: The evolution of creative ideas in engineering design course projects. *Design Studies*, *47*, 47–72. https://doi.org/10.1016/j.destud.2016.08.003

Toh, C. A., & Miller, S. R. (2014). *The Role of Individual Risk Attitudes on the Selection of Creative Concepts in Engineering Design*. ASME 2014 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, Buffalo, New York. https://doi.org/10.1115/DETC2014-35106

Toma, C., & Butera, F. (2009). Hidden Profiles and Concealed Information: Strategic Information Sharing and Use in Group Decision Making. *Personality and Social Psychology Bulletin*, *35*(6), 793–806. https://doi.org/10.1177/0146167209333176

Turesky, E. F., Smith, C. D., & Turesky, T. K. (2020). A call to action for virtual team leaders: Practitioner perspectives on trust, conflict and the need for organizational support. *Organization Management Journal*, *17*(4), 185–206. https://doi.org/10.1108/OMJ-09-2019-0798

Woolley, A. W., Aggarwal, I., & Malone, T. W. (2015). Collective intelligence and group performance. *Current Directions in Psychological Science*, *24*(6), 420–424. https://doi.org/10.1177/0963721415599543

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686–688.

Zheng, X., Ritter, S. C., & Miller, S. R. (2018). How Concept Selection Tools Impact the Development of Creative Ideas in Engineering Design Education. *Journal of Mechanical Design*, *140*(5), 052002. https://doi.org/10.1115/1.4039338

SUPPLEMENTAL MATERIALS

Screen Control

Screen control was accomplished by using the built-in functionality of Zoom. To facilitate screen control, the experimenter first uploaded a Line Rider track containing the course (Figure 2). Next, the experimenter shared their screen with the participants and told them that they would be working on the experimenter's computer during this phase using the screen takeover option in Zoom. Participants accomplished this by clicking "View Options" and then "Request Remote Control," after which the experimenter granted the participant access to the experimenter's screen. Due to limitations of the Zoom software, participants were not able to use the Command/Control Z option to undo their work, so they were instructed to ask the experimenter to do this for them if needed. Only one participant at a time could control the screen, so the participants were instructed to coordinate amongst themselves regarding who would control the screen and how they would pass off these responsibilities to other team members. Individual screen control time consisted of any directed movement (e.g., moving a mouse to select the pencil or paintbrush, load a file, point to things on the screen), drawing behaviors, erasing, adjusting the screen positioning, or other task-related behaviors requiring screen control.

Distribution of Influence Within Teams

Due to the constraints of the experiment, it is possible that track designs could share similar properties, even if they were created entirely independently (i.e., no influence). Thus, we expected some level of similarity between tracks due to chance alone. To determine the distribution of normalized difference values that should be expected due to chance alone, we conducted a permutation analysis (Puga-Gonzalez et al., 2021) using the designs produced in our study. First, we found the normalized differences between each team member's Brainstorming Phase track and the team's track in the Design Phase (Supplemental Figure 1). From these values, we also calculated their relative normalized differences by dividing the normalized differences of each team member by the sum of the normalized differences within their team (Supplemental Figure 2). Second, we permutated team designs across teams and repeated the above steps, but now comparing a given team's Brainstorming Phase tracks to a newly selected team track. We repeated this step 1000 times, generating 1000 permutated datasets. This approach holds consistent all observed aspects of designs produced in the Brainstorming and Design Phases but removes any potential effects of influence between the tracks produced in these phases. For each iteration, we recalculated the normalized differences and relative normalized differences between the original Brainstorming Phase tracks of each team and the newly selected team track.

In both the observed and permutated data, we sorted team members by their normalized difference values, placing them in the categories of "most influence" (i.e., lowest normalized difference), "intermediate influence" (i.e., intermediate normalized difference), and "least influence" (i.e. highest normalized difference). For the observed data, we calculated the median normalized difference and relative normalized difference values within each influence category. For the permutated datasets, we calculated the median normalized difference and relative normalized difference values within each influence category within each iteration. Our calculations for the permutated dataset resulted in a distribution of median normalized difference and relative normalized difference values that were expected due to chance. Within each influence category, we then compared the observed median normalized difference and relative normalized difference values to those expected due to chance. Observed values that fell outside the 95% confidence interval for these distributions were determined to be significantly different than expected (Supplemental Figure 3). We calculated p-values by finding the proportion of permutated values that had a level of influence that was greater than the value found in the observed data. If the proportion was greater than 0.5, we subtracted this value from 1. We then multiplied this proportion by 2 to obtain the two-tailed p-value (Farine, 2017).

We conducted a similar procedure to determine how much influence the Brainstorming Phase designs of participants in the Individual Condition impacted their designs in the Design Phase. We first calculated the normalized difference between each participant's design in the Brainstorming Phase and their design in the Design Phase (Supplemental Figure 1d). We then permutated Individual Condition Design Phase tracks across individuals. Next, we compared an individual's Brainstorming Phase track to the newly selected Design Phase track. We repeated this step 1000 times, generating 1000 permutated datasets. Within each iteration, we calculated the median level of normalized difference across all individuals. Finally, we compared the observed median level of normalized difference to the distribution of median normalized differences expected due to chance. Observed values that fell outside the 95% confidence interval for these distributions were determined to be significantly different than expected (Supplemental Figure 3b). P-values were calculated as described above.

Finally, we directly compared the observed level of normalized difference observed in individuals to that observed by teams' most, intermediate, and least influential team members using a Wilcoxon Rank Sum tests.

References

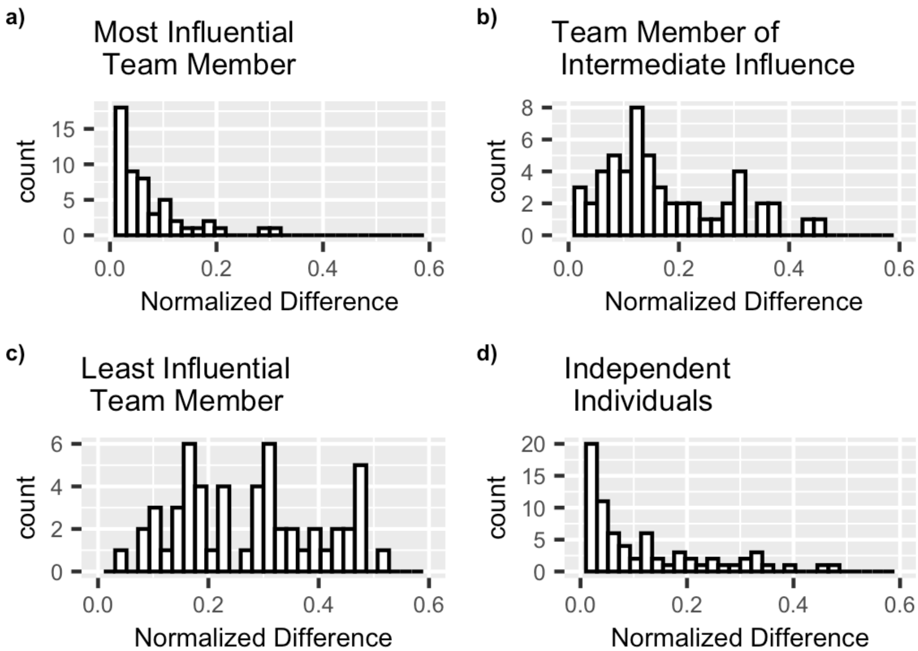
Farine, D. R. (2017). A guide to null models for animal social network analysis. *Methods in Ecology and Evolution*, *8*, 1309–1320. https://doi.org/10.1111/2041-210X.12772

Puga-Gonzalez, I., Sueur, C., & Sosa, S. (2021). Null models for animal social network analysis and data collected via focal sampling: Pre-network or node network permutation? *Methods in Ecology and Evolution*, *12*(1), 22–32. https://doi.org/10.1111/2041-210X.13400

SUPPLEMENTAL MATERIALS

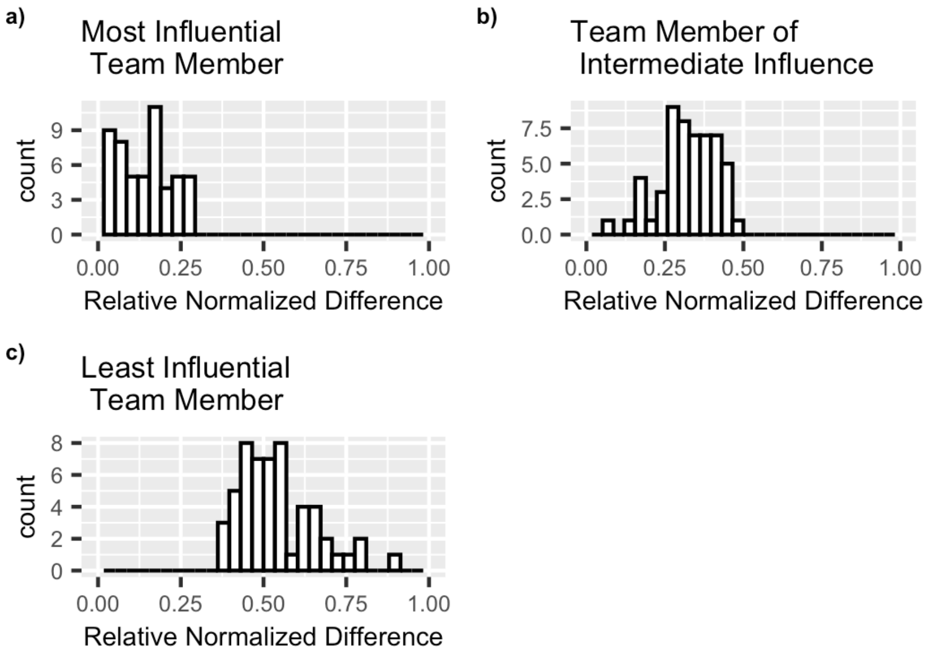
Supplemental Figures

Supplemental Figure 1



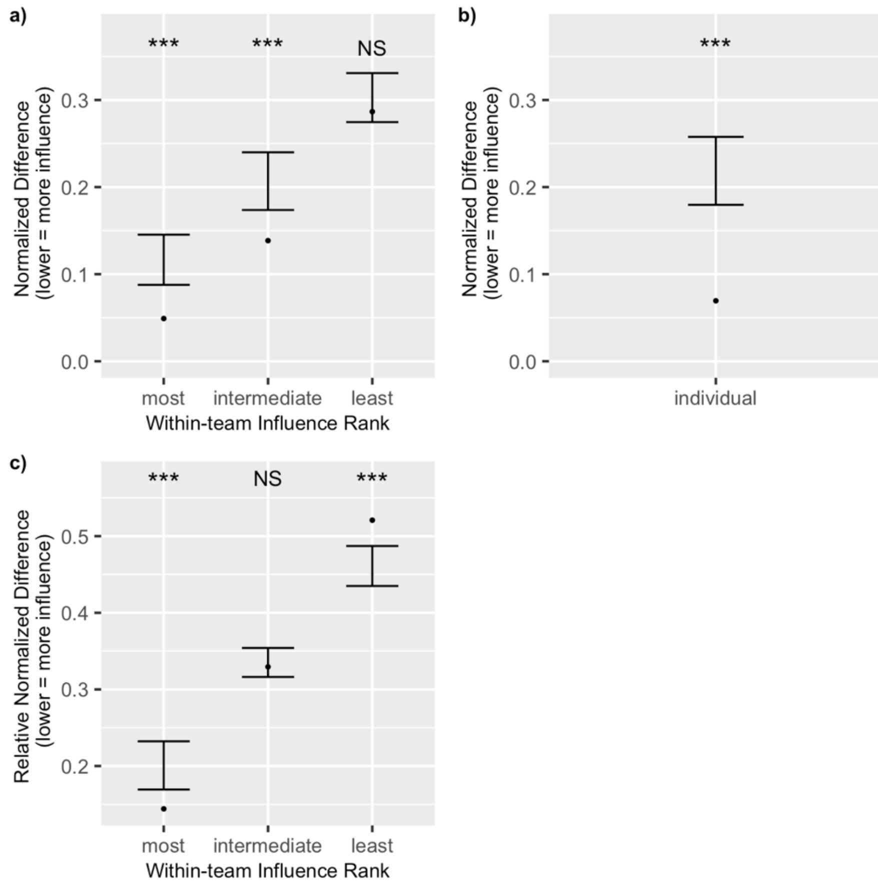
Note. Histograms of the normalized difference between a-c) team member Brainstorming Phase tracks and team Design Phase tracks and d) individual Brainstorming Phase and individual Design Phase tracks. Histograms for team members are divided according to whether team members were the a) most influential team member, b) the team member of intermediate influence, or the c) least influential team member. Lower normalized difference values indicate greater influence.

Supplemental Figure 2



Note. Histograms of the relative normalized difference between team member Brainstorming Phase tracks and team Design Phase tracks. Histograms are divided according to whether team members were the a) most influential team member, b) the team member of intermediate influence, or the c) least influential team member. Lower relative normalized difference values indicate greater relative influence.

Supplemental Figure 3



Note. Error bars represent the 95% confidence intervals for the levels of a,b) normalized difference or c) relative normalized difference expected due to chance alone. These distributions were generated by permutating Design Phase tracks across teams (Team Condition) or individuals (Individual Condition). The black points represent the median values observed across participants. Team members are divided into whether they had the most, intermediate, or least influence in the team, corresponding to the lowest, intermediate, and highest levels of normalized difference. *** indicates $p < 0.001$