



OPEN The anatomy of social dynamics in escape rooms

Rebeka O'Szabo^{1,2}✉, Sandeep Chowdhary³, David Derritei^{1,4,5} & Federico Battiston^{1b}✉

From sport and science production to everyday life, higher-level pursuits demand collaboration. Despite an increase in the number of data-driven studies on human behavior, the social dynamics of collaborative problem solving are still largely unexplored with network science and other computational and quantitative tools. Here we introduce escape rooms as a non-interventional and minimally biased social laboratory, which allows us to capture at a high resolution real-time communications in small project teams. Our analysis portrays a nuanced picture of different dimensions of social dynamics. We reveal how socio-demographic characteristics impact problem solving and the importance of prior relationships for enhanced interactions. We extract key conversation rules from motif analysis and discuss turn-usurping gendered behavior, a phenomenon particularly strong in male-dominated teams. We investigate the temporal evolution of signed and group interactions, finding that a minimum level of tense communication might be beneficial for collective problem solving, and revealing differences in the behavior of successful and failed teams. Our work unveils the innovative potential of escape rooms to study teams in their complexity, contributing to a deeper understanding of the micro-dynamics of collaborative team processes.

From the viral spread of rumours to the emergence of large-scale cooperation, human societies produce social dynamics and collective endeavours often hard to understand, characterize and predict. At the heart of this phenomenon is the innate need and ability of humans to collaborate and connect with others¹. Social interactions are indeed key to understand information exchange² and social contagion³⁻⁶. In recent years, advances in technologies have allowed us to track at unprecedented fine-grained scale face-to-face communication patterns in a variety of different contexts⁷⁻¹³, drawing attention to the importance of high-frequency time-resolved social processes¹⁴ to understand collective intelligence¹⁵ and collaborative problem solving^{11,16-18}. In the last decades, network science has proved to be a powerful and flexible framework to understand the complex relational structure of human dynamics^{19,20}, from structural balance theory^{21,22} to the detection of emergent mesoscale structures such as communities²³ and cores²⁴, associated with the coordinated behavior of multiple individuals in human societies.

From sport²⁵ to science²⁶, a crucial dimension of social interactions is encoded in teams. Indeed, whenever a goal needs to be achieved, people naturally join forces to overcome the limitations imposed by the capabilities of single individuals. In science, the importance of teams is increasingly recognised, as collective efforts are becoming more and more important for the production of knowledge, not only in volume but also in impact and attention²⁷. The predominance of team success has been attributed to the increasing need for specialized knowledge from different domains recombined through interdisciplinary collaborations to solve modern day problems²⁸⁻³⁰, highlighting the importance of team composition³¹. Past investigations about drivers of success in teams have revealed a complex and multifaceted picture. Diverse²⁶ and fresh teams³² may help widen skills and perspectives, though potentially introduce conflict and hinder efficient communication³³. Other known determinants of team success include collaborations across multiple institutions²⁹, inter-member familiarity³⁴ and prior shared successes of teams^{35,36}. Going beyond the simple dichotomy between solo and synergistic work, also size was found to be relevant, with larger teams developing science by focusing on more recent trends, in contrast to smaller ones, tending instead to dig deeper into past literature and often producing more innovative recombination of ideas³⁷. Even in apparently non-collaborative performances, such as those of elite athletes in individual disciplines such as tennis or martial arts, the presence of an organized, task-focused supportive team of people has become predominant and key for succeeding. All in all, teams serve as 'petri-dishes' for social influence and are often considered factories of innovation^{38,39}.

¹Department of Network and Data Science, Central European University, Budapest 1051, Hungary. ²Laboratory for Networks, Innovation and Technology, Corvinus University, Budapest 1093, Hungary. ³Department of Network and Data Science, Central European University, Vienna 1100, Austria. ⁴Department of Molecular Biology, Semmelweis University, Budapest 1085, Hungary. ⁵Channing Division of Network Medicine, Harvard Medical School, Boston 02115, USA. ✉email: rebeka.o.szabo@uni-corvinus.hu; battistonf@ceu.edu

In this work we study team dynamics in the innovative settings of escape rooms. Escape rooms are entertainment facilities that can be described as live-action board games, where a group of players has to collaborate intensely (search for clues, open locks, decipher codes) under time pressure to achieve the goal of the game by exiting the room within a predefined amount of time (typically one hour). Escape rooms are free from the typical frailties of traditional laboratory experiments and field studies. Similarly to experiments, they provide the same controlled environment for all groups under observation, yet minimizing the potential modification of participants' behavior as a response to being examined by researchers⁴⁰. Besides, social interactions can be followed and recorded at a high-frequency, allowing us to observe intact, non-manipulated teams in a nuanced, meticulous manner, overcoming the limitations associated with temporally aggregated data^{41,42} or self-reported questionnaires. Exploiting this innovative setting, we extracted from video records the real-time verbal and nonverbal communication of 40 small problem-solving teams. We analyze their high-resolution social dynamics, including collaborative network evolution and conversation rules guiding communication, exchanges of emotions and group interactions, linking them to successful team performance. Besides, we integrate such information by exploring the wider sociodemographic characteristics of team players, including gender, age and education, and their prior acquaintanceship, and meeting frequency. By capturing high-frequency real-time social interactions, our research highlights the potential of escape rooms to investigate task-performing teams in a minimally biased environment, contributing to advance the new science of teams.

Results

We study the social interactions of 40 teams collected in two distinct escape rooms. For each team (composed of 4 or 5 individuals, for a total of 171 players, all from Hungary) we extract high-resolution temporal interactions from video records (see “Methods”). Teams that manage to escape within one hour are deemed ‘successful’, while teams unable to do so are labeled as ‘failed’ groups. Teams must explore and exploit information, possibly through collaborations, by searching for clues, deciphering codes or opening locks. None of these tasks require any specific knowledge, skills or capabilities. However, experienced escape room players can be expected to have an advantage. Therefore, all teams in our data contain inexperienced members only (first-timers, or people who played maximum once in another room). For each team, we also recorded the sociodemographic characteristics of each player (such as age, gender and education), and relational data among team members (such as prior acquaintanceship and meeting frequency), by a questionnaire filled individually right before the game.

Our data can be mapped as a temporal network, where interactions among individuals occur at specific points in time⁴³. In Fig. 1 we present a schematic picture of collaborative social activities in escape rooms and its network representation at five different temporal snapshots. We record both verbal and non-verbal (e.g. showing something) interaction ties between team members, between one sender and one receiver (pairwise interaction) or more than one receiver (group interaction). Interactions are directed, and assigned either a neutral, positive or negative sign (see “Methods”). Summary statistics of the investigated groups are reported in the Supplementary Material. In the following, we provide a data-driven characterization of different dimensions of the social dynamics taking place in escape rooms.

General features of collaborative groups

Conversations in escape rooms are fast-paced, with teams having on average 30 interactions per minute among members, each one typically lasting 3 s. We find that older and more educated players speak for longer stretches at a time (Fig. 2A,B). The strength of a relationship between individuals is associated with the intensity of their

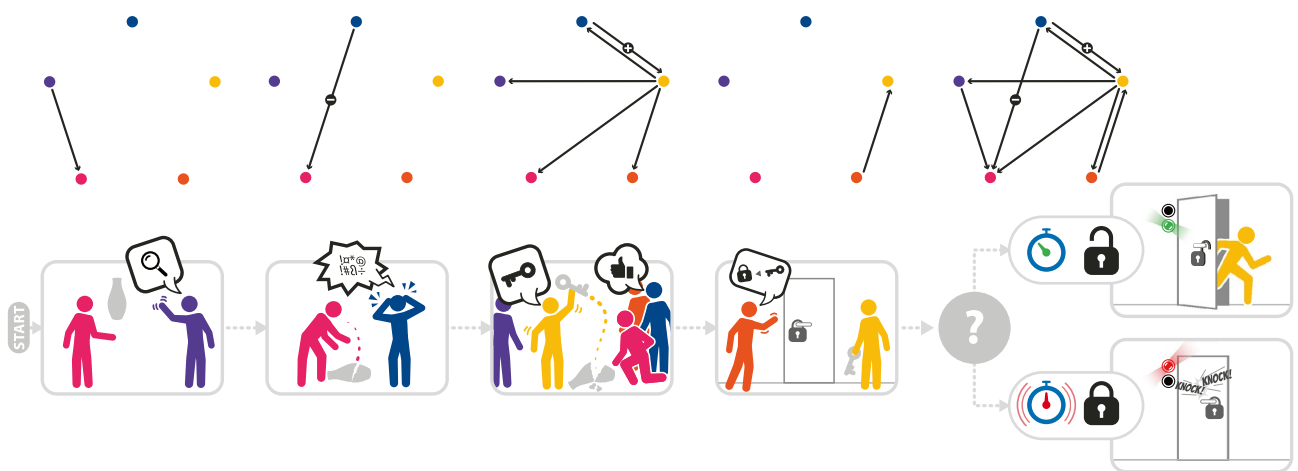


Figure 1. The complex and diverse dimensions of social interactions in escape rooms. We collect the social interactions of teams, who must collaborate in the non-routine environment of escape rooms to successfully solve tasks and thereby exit the room within an hour. Interactions are represented as temporal networks, which capture skeletal structure of communication between team members. Each interaction is directed, can be emotionally neutral or charged (either positively or negatively), addressed to a particular person (pairwise) or to a wider group. *This figure is produced by Szabolcs Tóth-Zs.

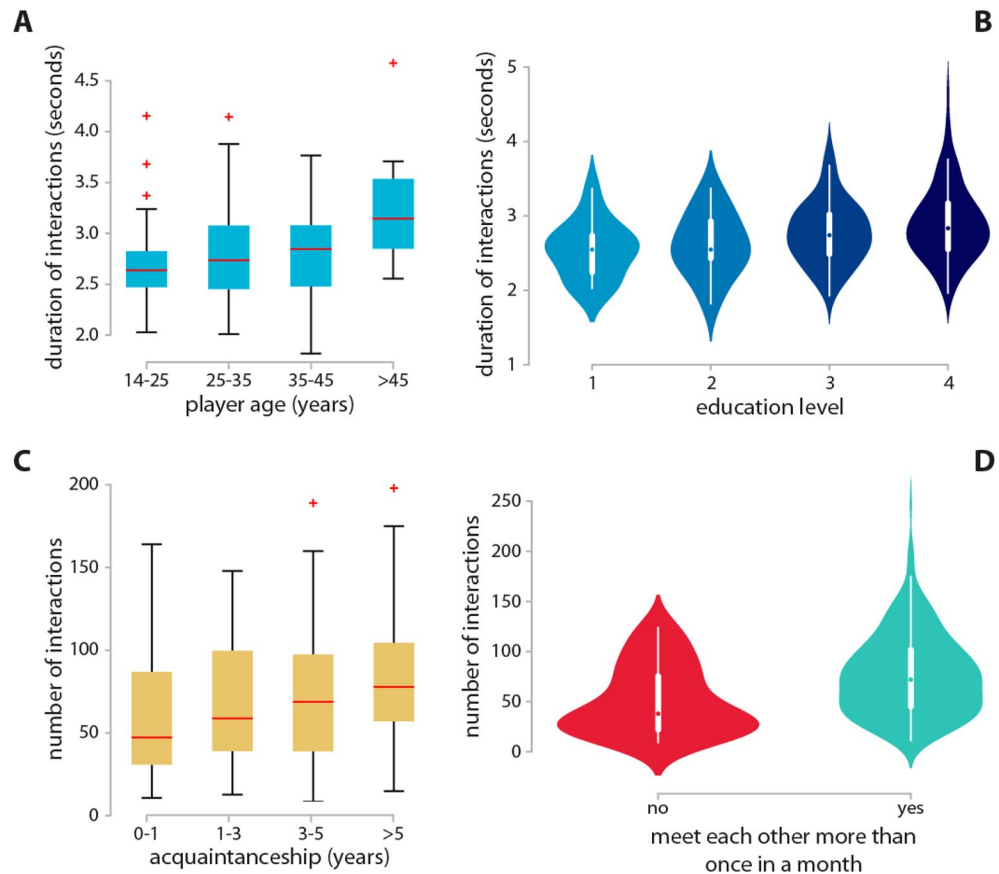


Figure 2. General features of team interactions. Individual demographics and relational characteristics of players are associated with their interaction patterns. Older (A) and more educated (B) players sustain longer interactions. Moreover, both the length of prior acquaintanceship (C) and meeting frequency (D) are correlated to an increased number of interactions during problem-solving.

collaborative pairwise interactions during problem-solving. In particular, prior member familiarity seems to promote communication, and the more time two players have known each other, the higher their rate of communications (Fig. 2C). Moreover, people who meet frequently (more than once in a month) interact approximately 1.6 times more during problem-solving than those with lower meeting frequency (Fig. 2D).

The microscopic architecture of team interactions

Conversations involve participants taking turns to speak, punctuated by stretches of silence. Changes in these turns among the participants are governed by rules, ensuring a basic level of order and intelligibility. To understand which rules govern collaborative communication in escape rooms, we use the Participation shift profile (P-shifts) framework developed by Gibson⁴⁴. P-shifts are behaviourally meaningful and easily interpretable inventory of network motifs, small subgroup patterns that carry information about the underlying mechanisms of social interactions. We count the frequency of all possible P-shifts associated with pairwise interactions, capturing the rules of conversation a team adopts. P-shifts can be categorized into three types—turn-receiving, turn-continuing and turn-usurping—for interpretation purposes (Fig. 3A), on which we elaborate in the following. By comparing motif counts with what is found in a suitable null model (see “Methods”), we determine whether the motif frequencies of the empirical data are significantly different from those observed by random chance in systems that preserve the same number of interactions, but where temporal correlations are washed away (Fig. 3B). We apply a 5 s upper bound as the maximum allowed timegap between successive interactions to be considered a valid P-shift (results do not depend significantly on the exact choice of this parameter).

The most frequent P-shift is *AB-BA*, a turn-receiving shift, where the second speaker *B* receives the entitlement to speak from the first speaker *A*. For this reason, the *AB-BA* motif is often referred to as the “current-select” rule, and usually covers questions, commands, requests or accusation-denial type of action-reaction pairs. The over-representation of this motif suggests a high demand for practical actions and task implementation under time pressure in escape rooms. Another turn-receiving motif, *AB-BY*, was instead significantly underrepresented in our data. *AB-BY* refers to those situations when an unaddressed recipient is turned into a target in the next speaking turn. This interaction chain is rare, in contrast to *AB-BA*, indicating that escape room players tend to have repeated exchanges as pairs, possibly associated with the presence of subgroups within teams. Turn

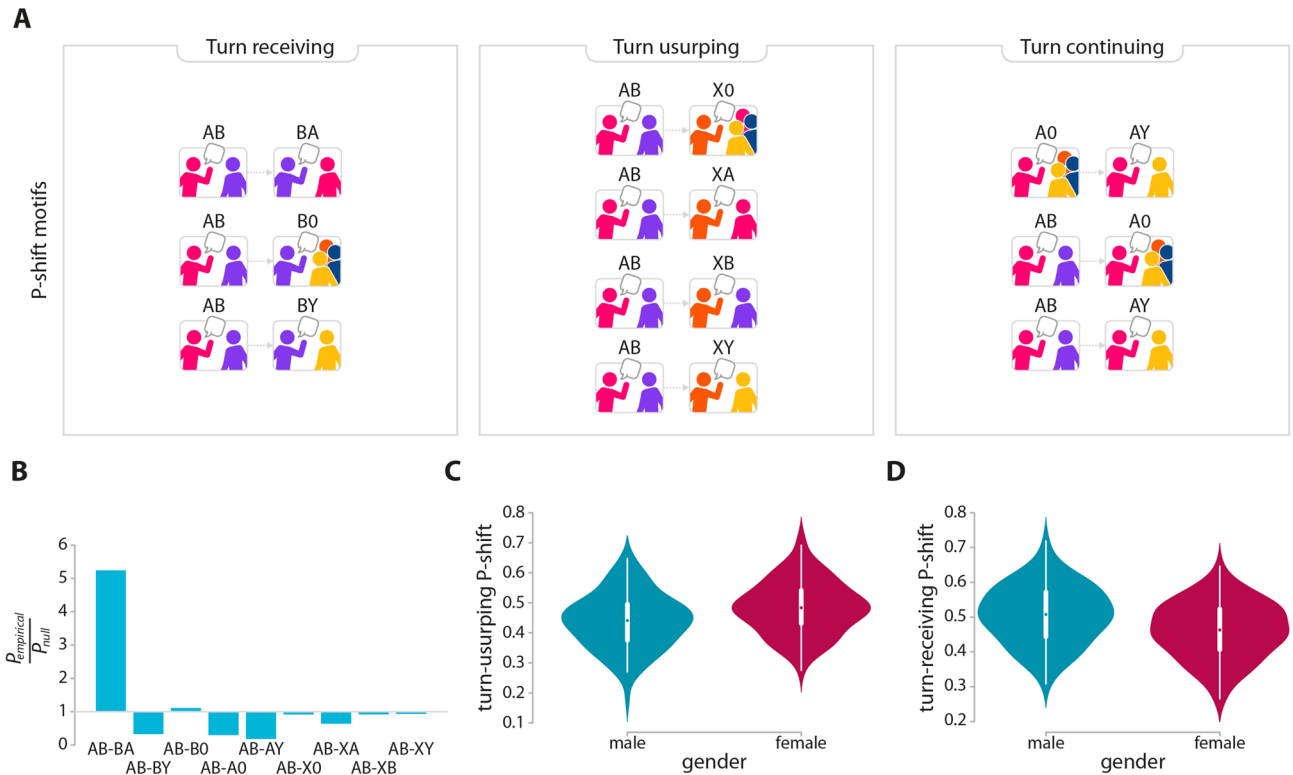


Figure 3. Microscopic nature of team interactions. We identify ordered communication sequences in escape rooms by Gibson’s Participation Shift Profile framework⁴⁴. P-shift motifs started with a directed remark (pairwise) can be clubbed up into 3 categories denoting the way by which the second speaker gets the turn (A). The prototypical turn-receiving P-shift *AB-BA* (which refers to an immediate reciprocation by *B* is overrepresented at the expense of *AB-BY*, while pairwise turn-continuing motifs such as *AB-AY* and *AB-A0* are underrepresented in escape rooms (B). Men and women participate differently in team conversations, with women showing lower levels of turn-receiving behaviour (C). This is probably connected to the observation that women are more likely to usurp turns (D) to express their opinion, as they rarely receive it. Panel (A) is produced by Szabolcs Tóth-Zs.

continuing motifs (such as *AB-A0* and *AB-AY*) where the same speaker shifts from one target to another, either a group or a third person, are also found to be statistically underrepresented.

Interestingly, a few roles tend to be associated with particular socio-demographic characteristics, such as age, education and gender, highlighting—in Gibson’s words—a *role differentiation*⁴⁴. This can be easily quantified by computing the Spearman’s rank correlation r_s and the Kolmogorov–Smirnov 2-sample test statistic D between a specific P-shifts motif and such features (see “Methods” for details). We find that older people are more inclined to address the group regardless whether they receive the turn (as *B* in *AB-B0*, $r_s = 0.17$, $p_{val} = 0.025$), or usurp the turn (as *X* in *AB-X0*, $r_s = 0.2$, $p_{val} = 0.025$). Players with the lowest level of education (elementary school) are more inclined to assume turn-usurper roles (both as *X* in *AB-XA*, $r_s = -0.2$, $p_{val} = 0.007$, and in *AB-XB*, $r_s = -0.16$, $p_{val} = 0.037$), though they are less likely to turn-usurp by addressing the group (*AB-X0*, $r_s = 0.16$, $p_{val} = 0.031$).

We also observe correlations between gender and given roles in escape rooms. We find that women typically turn usurp (as *X* in *AB-X0/AB-XA/AB-XB/AB-XY*, $D = 0.23$, $p_{val} = 0.02$, Fig. 3C) at the expense of receiving the turn to speak (as *B* in *AB-BA/AB-B0/AB-BY*, $D = 0.23$, $p_{val} = 0.019$, Fig. 3D). This shows that women are rarely given the opportunity to speak after being addressed, and that they have to rather seize the floor in order to express their opinion. Indeed, these associations are even more pronounced ($D = 0.61$, $p_{val} = 0.039$ and $D = 0.78$, $p_{val} = 0.003$, respectively) when we reduced our analysis to male-dominated teams only, where there is a single woman in the team. We also observe that the lack of turn receiving opportunities lead women in male-dominated groups to communicate much less than their teammates (as measured by their fraction of outgoing communication ties, $D = 0.611$, $p_{val} = 0.04$), a finding which is neither present in balanced groups nor in women dominated groups. This suggests that the social environment can impose constraints on women communication and relationship behavior in task-focused groups, complementing previous observations⁴⁵.

Signed interactions

Our experiment also gave us access to information about the emotional load of each interaction, classified as neutral, positive or negative. Praising, encouragement, and in general, every relationship-oriented behavior that has a positive effect on team spirit was classified as a positive tie. Negative edges were considered those that

influence team spirit adversely such as creating tension, provocation or disparagement. The wide majority of the recorded task-related exchanges were classified as neutral. However, if a task-related remark is emotionally loaded (e.g. somebody is yelling), a non-neutral sign is assigned to the communication tie (negative). This classification aims to mimic the well-established distinction between task-related and relationship-related social behavior^{46–48}. More details can be found in the “Methods” section.

We find that both older (Fig. 4A) and more educated players tend to have a smaller amount of emotionally loaded interactions ($r_s = -0.27$, $p_{val} = 0.0003$ and $r_s = -0.18$, $p_{val} = 0.021$, respectively). By looking at their signs, we observe that the fraction of negative ties are the ones that account for the previous, negative correlations ($r_s = -0.3$, $p_{val} = 0.00007$ and $r_s = -0.23$, $p_{val} = 0.003$, for age and education level, respectively), while these two socio-demographic features are not associated with higher chances of initiating positive interactions. These associations suggest that older and more highly educated team members are less likely to engage in social behavior that has a negative effect on team spirit.

Although there is a similar amount of emotional interactions in both successful and failed teams, Fig. 4B suggests that these emotional interactions are not equally distributed. Successful teams are more balanced in terms of emotional charge ($D = 0.40$, $p_{val} = 0.078$). In other words, members of successful teams tend to initiate a similar amount of emotional interactions with one another, while in failed groups an emotional polarization emerges, where only one or two actors display a higher number of emotionally loaded behavior. An investigation of how emotional ties evolve across time during the task performance reveals different temporal patterns for successful and failed teams (Fig. 4C). During the initial stages of the game, failed teams show a higher rate of emotionally loaded interactions. However, over time the rate of non-neutral interactions becomes higher for successful teams.

We can gain more insights into these patterns by decomposing signed interactions into positive and negative ties (Fig. 4D). Surprisingly, the initially greater amount of emotionally loaded interactions in failed teams is due to a high rate of positive interactions (13%, doubling the number of successful teams), which, however rapidly declines within the first 15 min. This unexpected feature might reflect a lower focus (e.g. making jokes, laughing),

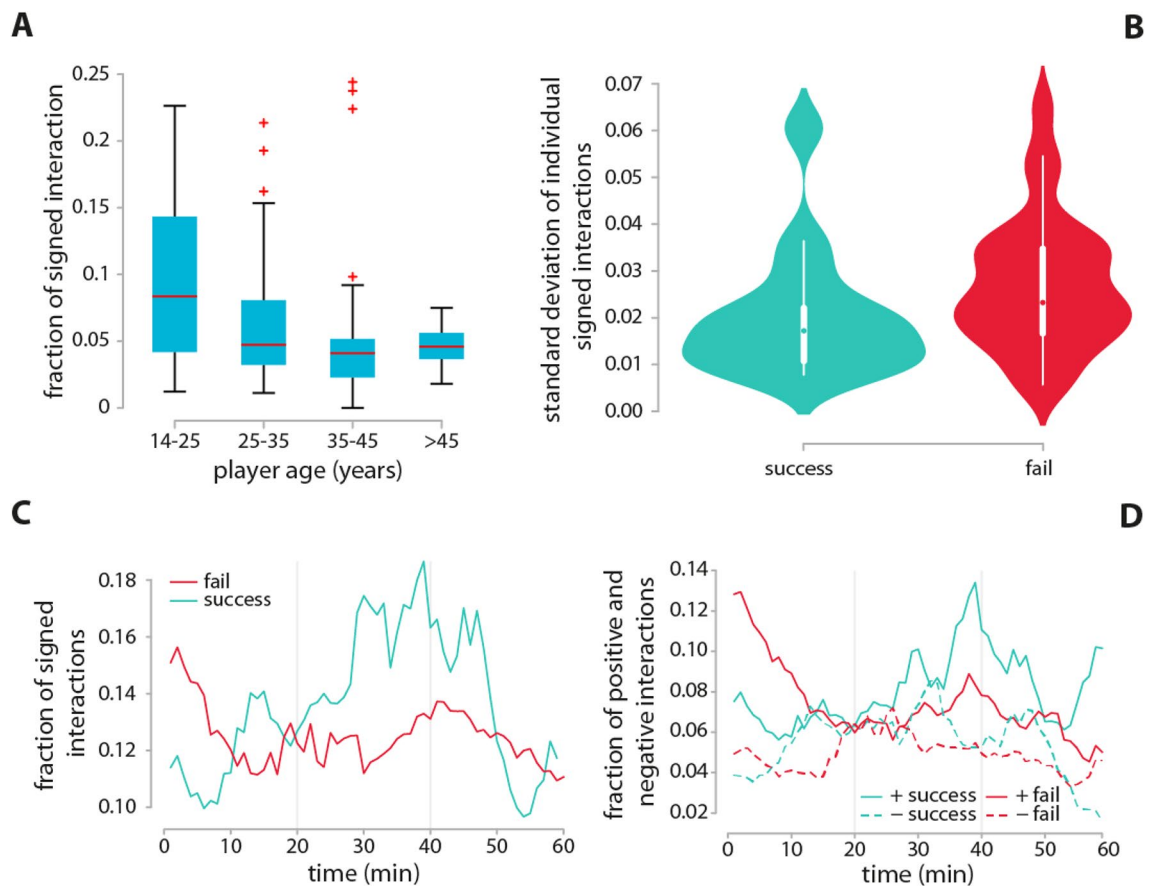


Figure 4. Signed interactions. Older players initiate less signed communication than younger members (A), suggesting a greater focus on task performance. While the total amount of emotionally charged interactions is similar in successful and failed teams, successful groups are more emotionally balanced across their individuals than the failed ones. This is quantified by the standard deviations of signed interactions across team members (B). Failed teams are characterized by a greater frequency of positive interactions in the initial stage of the game, which however rapidly declines over time (C). This trend is largely due to the high initial frequency of positive ties (D), possibly suggesting a lower amount of task focus. Surprisingly, successful teams show a slightly higher rate of negative interactions, suggesting that a minimum level of tense communication might be beneficial for collective problem solving.

which will eventually reveal crucial for the outcome of the game. As the game progresses, we find a greater frequency of positive interactions for successful teams, reaching 14% after 40 min, when many crucial tasks have already been completed. The dynamics of negative interactions is less rich. Negative ties are less frequent than positive ones for both successful and failed teams, and show no significant temporal patterns, with the exception of the rapid decline in negative interactions for successful teams towards the very end of the game. The slightly greater rate of negative ties in successful teams compared to failed ones (6% and 4% respectively) suggests that a minimum level of tense communication might be beneficial for task performance in collective problem solving. As some successful teams start leaving the room around the 47 min mark, only the remaining successful teams contribute to the curves shown in Fig. 4C,D.

Group interactions

Beyond pairwise communication, we also investigate group interactions⁴⁹, where more than a single recipient is addressed. Group interactions constitute $\approx 20\%$ of all the interactions (see Supplementary Material for more information), are equally divided into interactions involving 3 and 4 individuals together, and their duration is shorter, typically 77% of one-to-one interactions. We find that older (Fig. 5A) and more educated (Fig. 5B) people have a higher frequency of group interactions ($r_s = 0.23$, $p_{val} = 0.003$, and $r_s = 0.2$, $p_{val} = 0.009$, respectively), suggesting an unequal level of confidence and authority in the teams. Women display homophily when interact in groups (Fig. 5C), a phenomenon not observed when considering pairwise interactions only. No other sociodemographic data was found to be positively correlated with the presence of group interactions. All such analyses have been carried out by comparing the homophilic patterns observed in the real data with what is expected in corresponding null models where both group compositions as well as the number of interactions of each individual member is preserved.

Observing the temporal dynamics (Fig. 5D), teams tend to start the game with a high amount of group interactions. This pattern is associated with teams ‘warming up’, familiarizing themselves with the environment, and discussing together a strategy. Such a number falls abruptly after the first 5 min, reflecting division of work and individual focus. Such temporal group dynamics also display differences between successful and failed teams. Failed teams show a greater amount of group interactions around the 20-minute mark, associated with the emergence of potential early problems. By contrast, successful teams only display a low rate of group interactions around the same time, possibly reflecting greater focus and productivity. After manual visual inspection from the videos, we were able to associate the peak in group interactions in successful teams at the 30-minute mark to the presence of productive halftime get-together, where members synthesize their knowledge, discuss

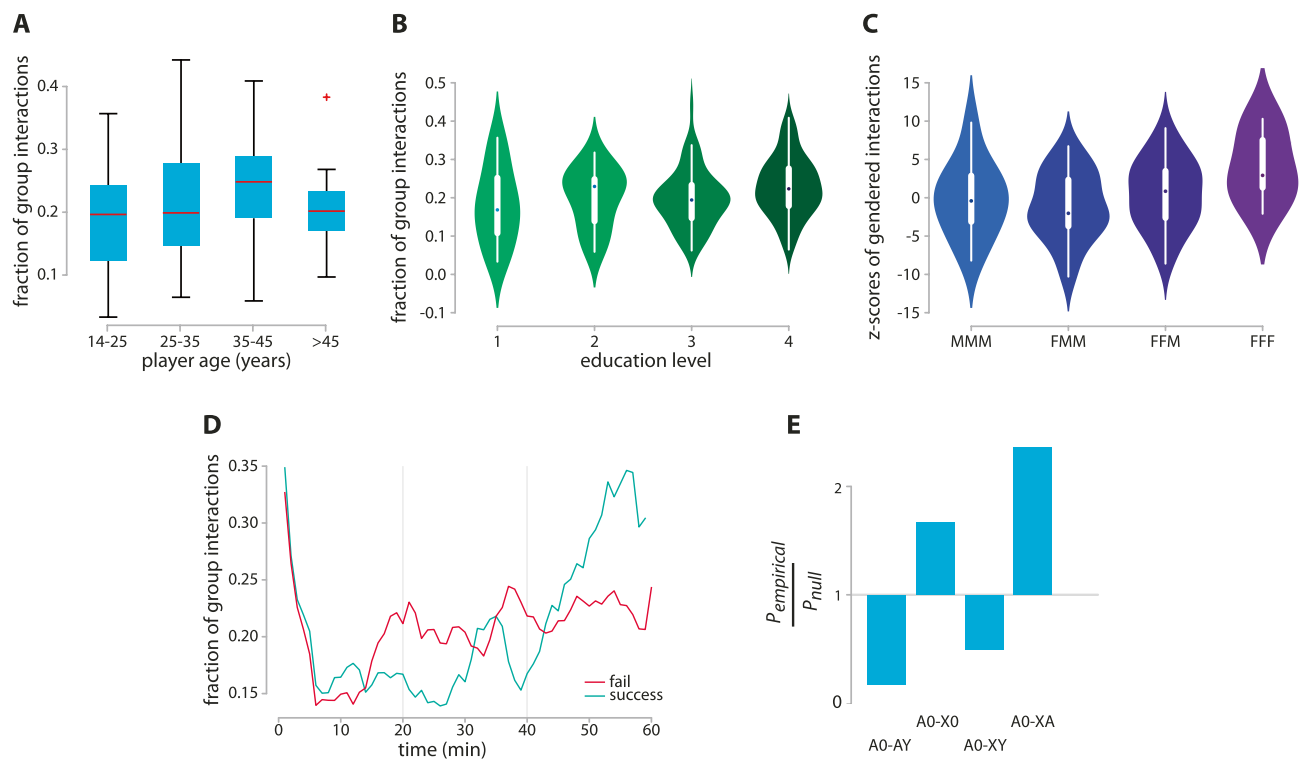


Figure 5. Group interactions. Older (A) and more educated (B) players display a higher frequency of group interactions, suggesting an unequal level of confidence and authority in the teams. Moreover, groups of women only tend to be overrepresented (C). An investigation of the dynamics of group interactions reveals that after an initial get-together members tend to work in smaller groups (D). An increase in group interactions for failed teams around the 20-min mark reflects the emergence of possible early issues, while the rapid increase for successful teams in the latest part of the game is associated with celebration. P-shift analysis reveals the overabundance of motifs A0-X0 and A0-XA, typically associated with the delivery of complex messages (E).

achievements how to proceed. As a few successful teams managed to escape slightly before the end of the game, only the remaining successful teams contribute to the curves shown in Fig. 5D.

Conversation rules governing group interactions can be analysed by a suitable extension of the previously introduced P-shift profiles⁴⁴. Following a group remark, where θ identifies the group, either the speaker continues their turn (as A in $A\theta$ - AY , turn continuing), or someone else claims the turn (turn claiming). Turn-claiming can happen in three distinct ways, as X in $A\theta$ - $XA/A\theta$ - $X\theta/A\theta$ - XY , where the addressee can be the first speaker A , the group θ , or another person Y . Observing the pattern of ordered group interactions accounting for conversation rules, we find that when someone claims the floor, it typically happens in the form of $A\theta$ - $X\theta$ or $A\theta$ - XA at the expense of $A\theta$ - XY (Fig. 5E). These two P-shifts, overrepresented in our data, cover those situations when a remark addressed to the group is followed by a reaction.

To assess the relevance of each motif, we compare its frequency in the real data against what observed by random chance in systems which preserve the same number of interactions but where temporal correlations have been eliminated (Fig. 5A, no significant difference between successful and failed teams). We find that the most overabundant motif is $A\theta$ - XA , the 'group version' of the previously discussed AB - BA pairwise motif⁴⁴. A second group motif, $A\theta$ - $X\theta$, is also found to be overabundant. Both of these P-shifts (someone addressing a group followed either by a second speaker addressing the group again or a team representative replying to the original communicator) are usually associated with the delivery of complex messages. For example, after A proposes an idea or gives an instruction to the group, X takes the floor to explain or translate A 's idea ($A\theta$ - $X\theta$), or asks for clarification to A 's action on behalf of the team ($A\theta$ - XA).

Discussion

Despite recent intensive efforts in characterizing human behavior from large-scale data, understanding the social and demographic drivers of successful team interactions is still a largely open and widely debated research area. Here we provided for the first time a characterization of the social dynamics of team interactions in escape rooms, non-interventional social laboratories previously unexplored in a fine-grained quantitative manner. By capturing high-frequency real-time social interactions in this innovative quasi-experimental setting, we were able to extract the building blocks of cooperative work. Our analysis revealed that socio-demographic characteristics might impact problem-solving communication. For instance, older and highly educated actors were observed to speak for longer, more often to the group, and initiate less emotional (in particular less negative) interactions, while prior strength of relationships between group members was positively associated with enhanced social interaction.

An investigation of P-shift profiles revealed the high demand for practical actions under time pressure, manifested in conversation rules such as the "current-select" rule (AB - BA), associated with frequent pairwise action-reaction exchanges. Interestingly, the behavior of men and women was found to be characterized by different conversation rules, with women often forced to turn-usurp in order to express their opinion, a pattern particularly overabundant in teams with only one female member. Successful teams displayed a higher emotional balance across their members, possibly reflecting the higher collective nature of team organization. In spite of a tendency for prosocial communication at the initial stage of the game, a temporal analysis revealed that already after 20 minutes failed teams had different interaction patterns from successful ones, displaying less task-focused behavior and the first signs of social conflict. Interestingly, also successful groups were found to maintain a non-negligible number of negative interactions until the end of the game, suggesting that a minimum level of tense communication might be beneficial for collective problem solving.

In summary, here we have proposed escape rooms as an innovative research setting for studying groups in a controlled, yet non-manipulated environment, where one can obtain high-resolution data on collective behavior. By investigating social dynamics at a fine-grained scale, we were able to portray an innovative and nuanced picture of the collective actions of these project teams. In the future, we intend to further investigate the division of work in these problem-solving groups to understand how exploration and exploitation tasks impose different demands on teams to allocate their resources, and how they are associated with bottlenecks slowing down collective performance. We hope that these insights will spark more research on team processes in escape rooms, easily accessible social laboratories contributing to a deeper understanding of intra-group dynamics and collective action.

Methods

Data processing

Interaction data on the problem-solving activity of teams were extracted from the video records by selected trained transcribers who were trained through several individual sessions. As a test of inter-coder reliability, following individual training, all applicants received the same video to code as a trial. In particular, consistency among and within transcribers was quantified through the Krippendorff alpha test, for which we obtained a value of 0.73, well above the conventional threshold of 0.65⁵⁰. After the transcribers selection was completed, randomly chosen samples of the transcriptions were double-checked in a systematic manner, obtaining similar consistency values. Transcribers were given a detailed guide on how to translate video information into a temporal signed edge list. Above all, transcribers had to consider that an interaction covers both verbal (one line of thought with maximum a few seconds of silence) and non-verbal communication (e.g., nodding). The start time and end time of an interaction is the second where the given communication is already/still present.

Signed interactions

We categorized the emotional load of the interactions and associated to each tie either a positive, negative or neutral sign. A precise guide on the classification of signed interactions was shared with the transcribers. In

particular, positive interactions are all interactions that relate to the previous action positively/have positive emotional charge and is not directed to the task itself but is rather related to interpersonal relationships. Typical examples of positive interactions are praising, laughing, joking and the encouragement/provision of emotional support of one another. In sum, all communication that are expected to influence the unity and spirit of the team positively is recorded as a positive interaction. In contrast, disparagement, despising, ironic, cynical reaction, provocation, intentional ignoring of someone, and in general, any other interaction creating tension between group members are considered negative. Therefore, negative interactions are those communication ties that relate to the previous action negatively and are not directed to the task, and thereby, they are assumed to influence team spirit and unity negatively. Task-related interactions, no matter if they express agreement or disagreement, are classified as neutral, as soon as they do not share any of the features described above, i.e. they do not carry emotional load. Typical neutral communication ties are task-related questions, replies and simple disclosure of facts. Taken together, neutral interactions cover all emotionally indifferent, task-focused communication, while signed interactions are those influencing team spirit positively/negatively, and which are constructive/disruptive to social relations in the context of a collaboration-demanding, pressed environment. In addition to this information, transcribers were given specific examples taken from the videos on managing correctly the categorization of signed interactions. Inter-coder consistency for signed interactions was evaluated as part of the consistency test described above. A 5-min rolling window was applied for plots showing the temporal evolution of signed interactions (Fig. 4C,D).

Group interactions

When three or more individuals were found to interact together, all actors involved were considered members of the same group, encoding communication beyond dyadic ties and traditional links. A detailed review of the formalism and mathematical tools used to properly encode group interactions can be found in Ref.⁴⁹. The ratio between the number of group interactions and the pairwise one can be computed by dividing the total number of hyperedges of order greater than two by the number of hyperedges of order equal to two. More information on the sociodemographic compositions of groups can be found in the Supplementary Material. A 5-min rolling window was applied for plots showing the temporal evolution of groups (Fig. 5D).

Null model for identifying P-shifts

To capture P-shift motifs observed in team communication patterns, we compare the abundance of each motif in the real data with what is expected in a suitable null model. In particular, we take the time series of interactions within a team during the play and obtain a randomized version of it by shuffling the order in which these interactions occur. This kind of randomization preserves the total interactions initiated/received by each individual, as well as the frequency and composition of group interactions. However, the temporal order of events is completely randomized. Any deviation from this null model signals the presence of temporal correlations in form of individual memory of past contacts, which do not merely depend on the activity and the total number of interactions of each individual.

Statistical tests

Spearman's correlation coefficient is a nonparametric measure of rank correlation. For this reason, it is able to assess the presence of positive monotonic relationships, whether they are linear or not. If there are no repeated data values, a perfect Spearman correlation of +1 or -1 occurs when each of the variables is a perfect monotone function of the other. The significance level of the found correlation is described by the corresponding p-value.

Kolmogorov–Smirnov (KS) 2-sample test is a nonparametric test typically used to assess whether two underlying one-dimensional probability distributions differ in a significant way. In particular, the KS test statistic is defined as $D_{m,n} = \max_x |F(x) - G(x)|$, where the two samples have sizes m , n , with observed cumulative distribution functions (CDFs) of $F(x)$, $G(x)$ respectively. The null hypothesis i.e. both samples come from the same distribution, can be rejected (at significance level α) if $D_{m,n} > D_{m,n,\alpha}$ where the critical value $D_{m,n,\alpha}$ can be found online in the corresponding standard table.

Ethical statements

The authors confirm that all experimental protocols were approved by the Ethical Research Committee of Central European University. We confirm that all methods were carried out in accordance with the Committee's Ethical Research Policy and Guidelines (https://acro.ceu.edu/sites/acro.ceu.edu/files/attachment/basicpage/141/ethic_alresearchpolicyandguidelines_0.pdf) including obtainment of informed consent from all participants.

Data availability

Data are available at: https://github.com/bambuszka/escape_room.

Received: 9 December 2021; Accepted: 30 May 2022

Published online: 22 June 2022

References

1. Uzzi, B. Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing. *Am. Sociol. Rev.* **20**, 481–505 (1999).
2. Granovetter, M. S. The strength of weak ties. *Am. J. Sociol.* **78**, 1360–1380 (1973).
3. Centola, D. & Macy, M. Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**, 702–734 (2007).
4. Aral, S. & Walker, D. Identifying influential and susceptible members of social networks. *Science* **337**, 337–341 (2012).

5. Lehmann, S. & Ahn, Y.-Y. *Complex Spreading Phenomena in Social Systems* (Springer, 2018).
6. Guilbeault, D., Becker, J. & Centola, D. Complex contagions: A decade in review. *Complex Spread. Phenomena Soc. Syst.* **20**, 3–25 (2018).
7. Pentland, A. Socially aware, computation and communication. *Computer* **38**, 33–40 (2005).
8. Eagle, N. & Pentland, A. S. Reality mining: Sensing complex social systems. *Pers. Ubiquit. Comput.* **10**, 255–268 (2006).
9. Pentland, A. & Heibeck, T. *Honest Signals* (MIT press, 2008).
10. Callahan, G. Pentland, Alex, social physics: How good ideas spread-the lessons from a new science, New York, NY: The penguin press, 2014. vii+ 320 pages. (hardback). *Rev. Austrian Econ.* **29**, 93–97 (2016).
11. Sekara, V., Stopczynski, A. & Lehmann, S. Fundamental structures of dynamic social networks. *Proc. Natl. Acad. Sci.* **113**, 9977–9982 (2016).
12. Lederman, O. *et al.* Open badges: A low-cost toolkit for measuring team communication and dynamics. [arXiv:1710.01842](https://arxiv.org/abs/1710.01842) (arXiv preprint) (2017).
13. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the Copenhagen networks study. *Sci. Data* **6**, 1–10 (2019).
14. Pentland, A. The new science of building great teams. *Harv. Bus. Rev.* **90**, 60–69 (2012).
15. Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N. & Malone, T. W. Evidence for a collective intelligence factor in the performance of human groups. *Science* **330**, 686–688 (2010).
16. De Montjoye, Y.-A., Stopczynski, A., Shmueli, E., Pentland, A. & Lehmann, S. The strength of the strongest ties in collaborative problem solving. *Sci. Rep.* **4**, 1–6 (2014).
17. Gomez, C. J. & Lazer, D. M. Clustering knowledge and dispersing abilities enhances collective problem solving in a network. *Nat. Commun.* **10**, 1–11 (2019).
18. Monechi, B., Pullano, G. & Loreto, V. Efficient team structures in an open-ended cooperative creativity experiment. *Proc. Natl. Acad. Sci.* **116**, 22088–22093 (2019).
19. Christakis, N. A. & Fowler, J. H. *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives* (Little, 2009).
20. Easley, D. *et al.* *Networks, Crowds, and Markets* Vol. 8 (Cambridge University Press, 2010).
21. Fritz, H. *et al.* The psychology of interpersonal relations (1958).
22. Rapoport, A. Mathematical models of social interaction. (1963).
23. Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proc. Natl. Acad. Sci.* **99**, 7821–7826 (2002).
24. Borgatti, S. P. & Everett, M. G. Models of core/periphery structures. *Soc. Netw.* **21**, 375–395 (2000).
25. Fewell, J. H., Armbruster, D., Ingraham, J., Petersen, A. & Waters, J. S. Basketball teams as strategic networks. *PLoS One* **7**, e47445 (2012).
26. Guimera, R., Uzzi, B., Spiro, J. & Amaral, L. A. N. Team assembly mechanisms determine collaboration network structure and team performance. *Science* **308**, 697–702 (2005).
27. Wuchty, S., Jones, B. F. & Uzzi, B. The increasing dominance of teams in production of knowledge. *Science* **316**, 1036–1039 (2007).
28. Hunter, L. & Leahey, E. Collaborative research in sociology: Trends and contributing factors. *Am. Sociol.* **39**, 290–306 (2008).
29. Jones, B. F., Wuchty, S. & Uzzi, B. Multi-university research teams: Shifting impact, geography, and stratification in science. *Science* **322**, 1259–1262 (2008).
30. Torrisi, S., Manfredi, S., Iacopini, I. & Latora, V. Creative connectivity project-a network based approach to understand correlations between interdisciplinary group dynamics and creative performance. *E & PDE 2019 Towards a New Innovation Landscape* 530–535 (2019).
31. Jackson, S. E., Joshi, A. & Erhardt, N. L. Recent research on team and organizational diversity: Swot analysis and implications. *J. Manag.* **29**, 801–830 (2003).
32. Zeng, A., Fan, Y., Di, Z., Wang, Y. & Havlin, S. Fresh teams are associated with original and multidisciplinary research. *Nat. Human Behav.* **20**, 1–9 (2021).
33. Uzzi, B., Wuchty, S., Spiro, J. & Jones, B. F. Scientific teams and networks change the face of knowledge creation. In *Networks in Social Policy Problems* 47–59 (Cambridge University Press, 2012).
34. Harrison, D. A., Mohammed, S., McGrath, J. E., Florey, A. T. & Vanderstoep, S. W. Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality. *Pers. Psychol.* **56**, 633–669 (2003).
35. Petersen, A. M. Quantifying the impact of weak, strong, and super ties in scientific careers. *Proc. Natl. Acad. Sci.* **112**, E4671–E4680 (2015).
36. Mukherjee, S., Huang, Y., Neidhardt, J., Uzzi, B. & Contractor, N. Prior shared success predicts victory in team competitions. *Nat. Hum. Behav.* **3**, 74–81 (2019).
37. Wu, L., Wang, D. & Evans, J. A. Large teams develop and small teams disrupt science and technology. *Nature* **566**, 378–382 (2019).
38. Börner, K. *et al.* A multi-level systems perspective for the science of team science. *Sci. Transl. Med.* **2**, 49 (2010).
39. Uzzi, B., Mukherjee, S., Stringer, M. & Jones, B. Atypical combinations and scientific impact. *Science* **342**, 468–472 (2013).
40. Landsberger, H. A. Hawthorne revisited: Management and the worker, its critics, and developments in human relations in industry. (1958).
41. Leenders, R. T. A., Contractor, N. S. & DeChurch, L. A. Once upon a time: Understanding team processes as relational event networks. *Organ. Psychol. Rev.* **6**, 92–115 (2016).
42. Kozlowski, S. W. & Klein, K. J. A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes. (2000).
43. Holme, P. & Saramäki, J. Temporal networks. *Phys. Rep.* **519**, 97–125 (2012).
44. Gibson, D. R. Participation shifts: Order and differentiation in group conversation. *Soc. Forces* **81**, 1335–1380 (2003).
45. Ni, C., Smith, E., Yuan, H., Larivière, V. & Sugimoto, C. R. The gendered nature of authorship. *Sci. Adv.* **7**, eabe4639 (2021).
46. Jehn, K. A. A qualitative analysis of conflict types and dimensions in organizational groups. *Admin. Sci. Q.* **20**, 530–557 (1997).
47. Simons, T. L. & Peterson, R. S. Task conflict and relationship conflict in top management teams: The pivotal role of intragroup trust. *J. Appl. Psychol.* **85**, 102 (2000).
48. De Dreu, C. K. & Weingart, L. R. Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis. *J. Appl. Psychol.* **88**, 741 (2003).
49. Battiston, F. *et al.* Networks beyond pairwise interactions: Structure and dynamics. *Phys. Rep.* **874**, 1–92 (2020).
50. Krippendorff, K. Reliability in content analysis: Some common misconceptions and recommendations. *Hum. Commun. Res.* **30**, 411–433 (2004).

Acknowledgements

R.O.S. and F.B. wish to thank to Balázs Lengyel and László Lőrincz for their feedback, and Szabolcs Tóth-Zs for his help producing the visualizations. R.O.S. also acknowledges useful conversations with Adam R. Pah. This publication is the result of research conducted for Central European University, Private University – CEU GmbH. It was made possible by the CEU Open Access Fund.

Author contributions

R.O.S. and F.B. designed research. R.O.S. designed the surveys, organized the experimental set-up and conducted the data collection. R.O.S. and S.C. analyzed the data, with support from D.D. R.O.S. and F.B. wrote the paper, with contributions from S.C. and D.D.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-13929-0>.

Correspondence and requests for materials should be addressed to R.O. or F.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022, corrected publication 2025