Problem-Solving Teams in Escape Rooms: Composition, Collaboration Networks, and Social Dynamics

Rebeka O. Szabó

Central European University
Budapest, Hungary
Supervisor: Prof. Federico Battiston

A Dissertation Submitted in Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in Network Science

2022

RESEARCHER DECLARATION

I Rebeka O. Szabó certify that I am the author of the work Problem-Solving Teams in Escape Rooms: Composition, Collaboration Networks, and Social Dynamics. I certify that this is solely my own original work, other than where I have clearly indicated, in this declaration, the contributions of others. I also confirm that this work is solely based on the data I have collected on my own. The thesis contains no materials accepted for any other degrees in any other institutions. The copyright of this work rests with its author. Quotation from it is permitted, provided that full acknowledgement is made. This work may not be reproduced without my prior written consent.

Statement of inclusion of joint work

I confirm that Chapter 2 is based on a paper which was written in collaboration with Júlia Koltai and Federico Battiston. I came up with the idea of expanding faultline research and contribute to the categorization-elaboration model with the data I collected. Prof. Koltai drew my attention that my data would make a perfect fit to a survival analysis and actively supervised the process of building and specifying the models. I carried out the analysis of the data and developed the paper. Both Prof. Koltai and Prof. Battiston made essential contributions in editing and finalizing the paper.

I confirm that Chapter 3 is based on a paper which was written in collaboration with Sandeep Chowdhary, Dávid Deritei and Federico Battiston. Beside my longstanding notion of studying the interaction pattern of teams and relate it to performance, all my collaborators contributed to the analysis with valuable ideas. I would like to highlight Mr. Chowdhary's devotion and active collaboration in the process of actual data analysis, and Prof. Battiston's inestimable guidance with writing and finalizing the paper.

Signature of PhD Candidate:

March, 2022

Signature of Dr. , endorsing statement of joint work:

March, 2022

Signature of Dr. , endorsing statement of joint work:

March, 2022

ABSTRACT

From sport to science, a crucial dimension of social interactions is encoded in teams, 'petri-dishes' for social influence and factories of innovation. 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. What are the determinants of successful collective problem solving, and how do the network of social interactions evolve during creative tasks performance? This dissertation provides a fresh perspective on the topic by investigating – for the first time in a quantitative manner – the problem solving behavior of teams in escape rooms, a new non-interventional and minimally biased social laboratory. Escape rooms are entertainment facilities where a group of players has to collaborate intensely under time pressure to fulfil the goal of the game and eventually exit the room. Escape rooms are free from the typical deficiencies of traditional laboratory experiments and field studies. They provide the same controlled environment for all groups under observation, yet without interventions. Moreover, they allow us to observe intact teams in a nuanced and meticulous manner, overcoming the limitations of self-reported questionnaires and poorly time-resolved data. Exploiting this innovative setting, I extract from video records the real-time verbal and nonverbal communication of 40 small problem-solving teams and integrate such information with the sociodemographic characteristics of team players. This unique fine grained data enables me both to enrich social scientific theory and to contribute to the development of a more sophisticated picture of collaborative team dynamics. First, I investigate the role of diversity in team success. I use faultline theory that argues that diversity research should consider the combined, cumulative effect of different diversity dimensions to better understand its negative influence on team outcomes. I extend faultline theory by studying the moderating effect of team familiarity and cross-subgroup communication – the latter one is operationalized in an innovative way as real-time interactions – in the diversity-performance relationship. The key finding indicates that group faultline affects team success negatively *when* the faultline-induced subgroups do not communicate enough with each other. Next, I explore the evo-

lution of teams' social dynamics. I use tools from network science to provide a time-resolved map of team interactions, extract key conversation rules from motif analysis, and discuss turn-usurping gendered behavior, a phenomenon that I find to be particularly strong in male dominated teams. Investigating the temporal evolution of signed and group interactions I show that a minimum level of tense communication might be beneficial for collective problem solving, and I also reveal differences in the group behavior of successful and failed teams. Taken together, this dissertation 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.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor Federico Battiston for his continued support and enabling help. His guidance has been invaluable and his patience is much appreciated. I can say I have been lucky to have him as a supervisor.

I would like to acknowledge several other people who have influenced this work. I thank my closest colleagues and collaborators Dávid Deritei, Sandeep Chowdhary for their useful contribution to my work. Srebrenka Letina provided me with inestimable emotional support. The rest of the Department of Network and Data Science, including Olga Peredi, the other students and visitors provided a memorable environment to grow both professionally and personally.

I am also grateful to my colleagues at the Laboratory for Networks, Technology and Innovation, Corvinus Univserity Budapest, László Lőrincz, Balázs Lengyel and the all the other amazing people; my host at Kellogg School of Management and the Northwestern Institute on Complex Systems, Adam R. Pah.

I want to thank my husband Szilárd, my family and friends for their unlimited encouragement and support.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Social life naturally organizes into groups that satisfy the innate need of humans to connect to others [2], and facilitates coping with various problems, especially those requiring collective action. Groups face with typically poorly structured problems all over human history (such as environmental pollution or food shortage), as they cannot purely act based on the expected costs and benefits [3, 4] due to social dependencies between members. Group members may have different views of the problem itself, on its solution, and on the desired amount of individual contributions and benefits [4]. When groups transcend the basics of these dilemmas by developing joint intentionality and focus of a common goal [5], they behave as teams. Teams, when efficient, turn groups into dynamic and complex problem solving systems. Thus, in our rapidly changing and increasingly complex world, from sport [6] to science [7], teams appear ubiquitous.

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 attention [8]. 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 problems [9, 10, 11], highlighting the importance of team composition [12]. Past investigations about drivers of success in teams have revealed a complex and multifaceted picture. Diverse [7] and fresh teams [13] may help widen skills and perspectives, though potentially introduce conflict and hinder an efficient communication [14]. Other known determinants of team success include collaborations across multiple institutions [10], inter-member familiarity [15] and prior shared successes of teams [16, 17]. 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 delve deeper into past literature and often producing more innovative recombination of ideas [18]. Even in apparently non-collaborative performances, such as those of elite athletes in individual disciplines such as tennis or martial arts, the presence of 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 [19, 20].

Among the various theoretical roots of the network perspective (such as the rational self paradigm [21] or the social exchange school [22]) applied to group research, the communication network approach proved to be the most optimal to account for the dynamic nature of team processes. From this perspective, a group as a network is built up from nodes indicating group members, however, ties, as opposed to representing static or long-standing relationships denote interpersonal interactions which are inherently dynamic and temporal. Early research conducted specifically on small group networks to measure the effect of various interaction patterns on group functioning and performance primarily built on experimental design. These studies examined artificial, randomly composed groups in laboratories where experimenters manipulated the group's communication patterns in numerous ways (e.g., who can send messages to whom) [23, 24, 25], and found that groups benefited from centralized structures (hubs of communication) when the task to be implemented was simple, but it meant a disadvantage when performing complex tasks. More current studies expanded the scope of the examination of the optimal network structure for group performance by moving from the laboratory to the field [26]. Field studies on intact (already existing) teams concluded that actively communicating groups outperform groups with fewer ties [27, 28]. However, interaction data gained by self-reports in these cases appeared as a proxy of relational strength [29, 2], rather than as ties covering real interactions between team members.

This dissertation proposes a new, innovative setting for studying teams in their complexity: escape rooms. Escape rooms are free from the typical weaknesses 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 researchers [30]. Another advantage is that 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 [31, 32] 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. In addition, we obtained information about the wider sociodemographic characteristics of team players, including gender, age and education. Moreover, we gathered data on prior group relations such as acquaintanceship, and meeting frequency, as well as on leadership. First, leveraging our unique data, we address a predominant question in the field of organization and team science, the impact of diversity on performance. Then, we analyze teams' high-resolution social dynamics, including collaborative network evolution, conversation rules guiding communication, exchanges of emotions and group interactions, exploring their influence on successful team performance.

This thesis highlights the potential of escape rooms to investigate task-performing teams in a minimally biased environment. Moreover, it contributes to the increasingly rare studies of interacting groups that carry out actual activity as opposed to merely thinking about something. In addition to these advantages of this setting, escape rooms also present a cost-efficient research environment for capturing high-frequency real-time social interactions, a relatively sparse datatype due to the time-effort-and resource-intensive nature of collecting such information. However, teams cannot be understood in their complexity without studying social interactions, a core process that constitutes this complexity. Thus, this dissertation presents an original way to advance the new science of teams as complex systems.

We proceed with the thesis as follows:

- Chapter 2: Here we introduce escape rooms as non-interventional social laboratories suitable for studying collaborative problem-solving behavior at a high time resolution. We discuss the strengths of escape rooms over laboratory experiments and field studies. We present our fine-grained data and discuss the methodological considerations and the process of its collection.

- Chapter 3: In this chapter we extend diversity faultline research by investigating the predictors and moderators of team success. In particular, we investigate two moderators in the diversity-performance relationship that have not been examined before this way: team familiarity and the communication between the faultline induced subgroups that we expect to modify the negative effect of faultline in team success. We test our hypotheses by employing survival models, and demonstrate that team diversity can only be understood in relation to team processes. In particular, we show that teams with strong faultlines tend to fail *when* the faultline-induced subgroups do not communicate enough with each other.

- Chapter 4: In this chapter we present a data-driven, exploratory case study where we truly utilize the high time resolution of our data to reveal social behavior in escape rooms. The analysis portrays a nuanced picture capturing different dimensions of the social dynamics within teams including – collaborative network evolution, conversation rules guiding communication, macroscopic features of project teams, and patterns of successful task performance. Among the various identified characteristics of problem solving team dynamics, we reveal a phenomenon of turn-usurping gendered behavior and find a peculiar emotional dynamics where a minimum amount of tense communication seems to enhance team success.

We conclude by linking and highlighting our main findings, and discussing some of their potential interpretations. We also suggest future research directions, and present the preliminary results of an ongoing work where we zoom into the data and investigate efficient problem solving at the level of the subtasks rather than the level of teams.

CHAPTER 2

ESCAPE ROOMS: AN INNOVATIVE EXPERIMENTAL SETTING

2.1 Prelude

Recognizing that teams are complex systems, and aiming to understand them, it is necessary to investigate the core processes that maintains that complexity. Communication is a social means, which teams can exploit to meet effective functioning and performance [33]. Thus, interactions is the very heart of any team process. There are two predominant methods to investigate teams' interpersonal interactions: by laboratory experiments and field studies.

Laboratory experiments are characterized by artificially constructed research settings, laboratories, where researchers examine research subjects. Research subjects are explicitly recruited to participate in the given study, and most often the potential range of their behavior is constrained and instructed by the researchers (see studies of Group Networks Laboratory at MIT [23, 24, 25], where researchers manipulated the patterns of group communication by defining allowed and prohibited communication channels between the members of the team, and measured the effect of different communication patterns of group functioning and performance). Due to the pronounced restraints on who can interact with whom imposed by the experimenters, traditional laboratory research could address a limited scope of research questions [34].
Besides, a typical critique of traditional experiments on group interactions is that they are exclusively preoccupied with the communication networks of ad hoc, zero-history groups [35]. Therefore, they lack the attention to the embedding context of teams such as pre-existing relations between team members and their sociodemographic characteristics. Consequently, these studies incline to

treat groups as universal, generic entities [35], as if all teams were the same regardless of their characteristics.

The aforementioned weaknesses of traditional experiments might have contributed to the shift of many small group and communication researchers' interest from the laboratory to the field in the mid 1970s [36]. In this era, network research flourished in the field of sociology, and small group networks became an independent research area [34].
Broader in their perspective, field studies do take group characteristics into account while investigating group interactions. However, most of these studies either measure interaction process by retrospective questionnaires, and/or rely on an input-output design, thus focus on the two endpoints of group activity. Thereby they tend to overlook tremendous meaningful data on actual ongoing processes and links that they originally meant to observe.
Retrospective questionnaires collect group members' ratings or reports about what happened in the team previously. As a consequence, information yielded by retrospective questionnaires is necessarily coarse-grained data. Moreover, and most importantly, it gives entirely subjective information that is likely to include biases, such as highlighting the earliest or most salient interaction (e.g., interaction with a colleague) at the expense of all interactions [37, 38].
Furthermore, self-reports of team members are likely to be influenced by prior group life, making it implausible to provide an objective summary on what happened on a certain group occasion. Consequently, in spite of its widespread application, retrospective questionnaire is an inappropriate choice to collect fine-grained interaction data. Similarly, the input-output research design, is inadequate in capturing the dynamic nature of group interactions, since it does not intend to obtain data on actual interactions, but measures and manipulates the input and output variables, while producing inferences about what is going on in between [35].

At the same time, we cannot go by without mentioning the technological development of the recent years, that has allowed us to track at unprecedented fine-grained scale face-to-face communication patterns in a variety of different contexts from collective intelligence [39] to collaborative problem solving [40, 41, 42, 43]. Highlighting a few relevant, higher impact research, the Copenhagen Networks Study overcame the difficulties of collecting high resolution interaction data relying on smartphones by which Lehman et al. measured the networks of physical proximity, phone calls, and text messages and contributed to our understanding on human behavior with a sophisticated picture about social dynamics [41, 37]. Another promising endeavor acquiring temporal interaction data was formed in the Human Dynamics Group at MIT Media Lab. Pentland et al. developed wearable devices called Sociomet-

ric Badges with the aim to automatically quantify face-to-face social behavior [44, 45, 46, 47], and found that communication patterns are strong predictors of team performance [48]. These studies utilizing recent technological advances mean great progress in data collection compared to more traditional ways of investigating group dynamics and collaboration networks, such as laboratory experiments or field studies using self-reported data. Yet, laboratory experiments have the strength of observing groups in a controlled environment, and thereby minimize the occurrence of confounding factors.

Here, I present an innovative research setting for collecting fine-grained data on collaborative problem-solving teams. In particular, I introduce escape rooms as non-interventional social laboratories that combine the advantages of both laboratory experiments and field studies. At the same time, they are free from the typical weaknesses of these research methods. In the following section, I provide a detailed description of this research design, including the data collection and processing.

2.2 A quasi-experimental setting

Escape rooms are entertainment facilities that can be described as live-action board games. There is a voluntary group of people wishing to play (e.g. colleagues for team building, families, friends, etc.) who book an appointment and a room reflecting the atmospheric and cultural theme of the game (e.g. room Godfather, Fig. 2.1). After a short briefing about the rules, they get locked into this room. The goal is well-defined and understandable to every participant: they have to exit the place within a one-hour time frame. In order to do so, group members must collaborate by searching for clues, solving puzzles, opening locks, deciphering codes, etc. None of these tasks require prior preparation, any specific knowledge, skills, or capabilities. Nevertheless, the whole process of the game demands a high level of cooperation due to the time pressure, the unfamiliar nature of the environment, and the lack of any additional information describing or specifying the tasks in the room. Thus, group members must rely on each other, and thereby they behave as a team while members become interdependent by sharing a collective goal.
Moreover, time pressure under these collaborative groups have to explore and communicate in order to find a solution makes the research field circumstances analogous to the organizational settings of project teams performing non-routine assignments. In fact, escape room participants can be considered as actual project teams, as members are interdependent along a clearly-defined focus of a collaboration-demanding, non-routine task [49].

From a sociological perspective, escape rooms are social laboratories, as all

group processes take place in the same, above-detailed, controlled environment for all participating groups. At the same time, escape rooms are free from the typical weaknesses of laboratory experiments, as they are not artificially constructed mediums for conducting research. Moreover, the teams I study in escape rooms are intact, non-manipulated groups that are not torn out of their embedding social context, but "real" teams with pre-existing social relations observed in a noninterventional, yet controlled environment. Furthermore, unlike conventional field studies, I gather fine-grained and objective data about real-time collaborative interactions of these problem-solving teams in contrast with collecting self-reports about what has happened in the team previously.

2.3 Data

2.3.1 Collection

Data collection in escape rooms happened by video recording the whole problem-solving activity of collaborating groups. At the escape room company, (Paniq rooms) I went to gather data, game activity can be constantly observed through three cameras (recording the scenes from different angles), which originally serve to prevent inadequate acts (e.g.: destruction of objects in the heat of the game). Besides the general GDPR contract of the escape room, I asked the participants to sign an additional consent form (see Appendices 6.0.2), in which they agreed that the video recording of their play will be analyzed in this research. As video recordings are routinely acquired in these escape rooms, and the purpose of the participants is to participate in the game, and not to take part in the experiment, the condition of our access to the recorded data is not likely to modify participants' behavior. As a consequence, our data collection minimizes the Hawthorne effect [30, 50] - another common weakness of laboratory experiments - as it does not elicit the perception of being watched and expected to behave in specific ways that favor researchers' anticipation. Therefore, the teams examined are intact and non-manipulated.
My research sample consists of 40 teams where all members are inexperienced: first-timers or members who have had a maximum of one prior exposure to a different escape room in the past. As mentioned above, none of the tasks require any specific knowledge or capabilities. Therefore, all inexperienced escape room players can be considered equally competent in problem-solving. However, it is reasonable to envisage people familiar with escape room games to have comparative advantage over first-timers. Therefore, excluding teams with knowledgeable members in terms of escape rooms aimed to ensure the homogeneity of the groups and their comparability.

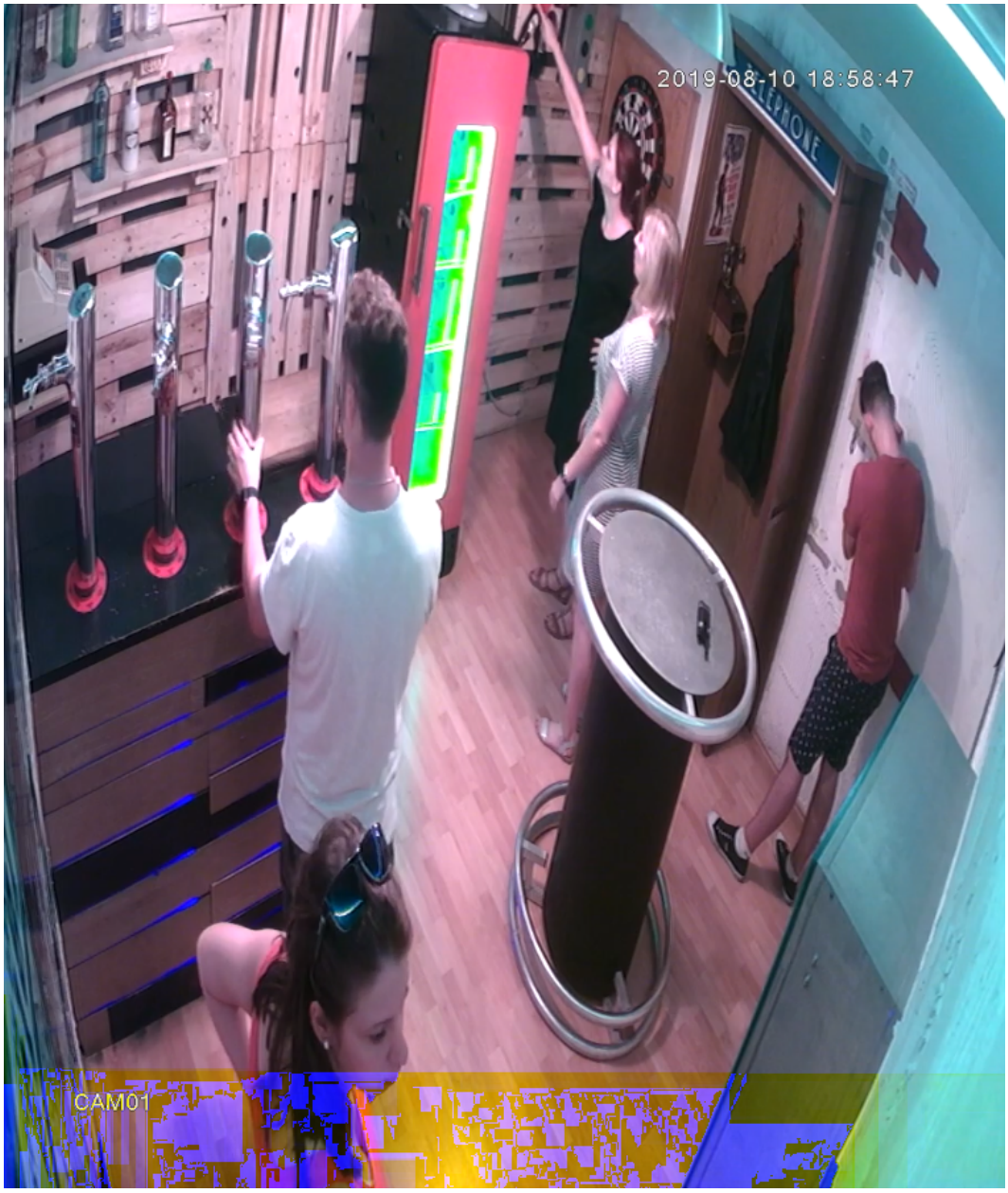


Figure 2.1. *A snapshot of problem solving in room Godfather.*

In selecting the two rooms for our experiment, an important consideration was that the gameplay should mimic a project in organizational environments as much as possible. I opted for two rooms where the complexity and the non-linear nature of the gameplay resemble team-based projects in organizations.

I refer to the two rooms as Sherlock and Godfather reflecting the atmospheric and cultural themes of the games. The tasks are nonlinear as there is a hierarchy of subtasks that eventually lead to the final solution of escape. The essence of nonlinearity in this sense is the fact that team members can work on different tasks in parallel, but occasionally have to combine solutions to move forward. The escape rooms Sherlock and Godfather are accessible for 2-6 people, however, they were originally designed for groups of 4-5 members. For this reason, I limited my investigation to teams with this intermediate size to avoid excessive variation in the data and the potentially spurious effects of team size in my analysis. The examined 40 teams are composed of 171 Hungarian participants with an average team size of 4.275.

As mentioned before, the teams are intact and non-manipulated. The game is suited to participants from the age of 12, however, such young players might lack associative capabilities and explorative skills compared to older ones. Young children are also likely to form a team with and overly rely upon their parents, which entails specific intra-group relations that are outside the scope of my study. Hence, my research sample covers groups with members who are at least high school students and assumed to be self-sufficient in problem-solving. Besides the video records, I designed a questionnaire for making inquiries on the social context the studied teams are embedded in (see in Appendices 6.0.3). Particularly, I collected data about the social-demographic characteristics and the prior social structure of the groups. In the former category, I asked participants about their age, gender, and highest level of education, while in the latter one, I was interested in how long team members have known each other and how frequently they meet. In addition, team members were asked to name the leader of their team, if any. The survey was filled out prior to problem-solving.

2.3.2 Processing

Once I obtained the videos on the project teams' problem-solving activity, I had to extract the interaction data. I registered all interpersonal interactions that occurred during task performance into an edgelist (Fig. 2.2) from which I constructed the network of collaborative communication. In particular, I recorded the following parameters of problem-solving interactions: the sender/source and the receiver/target of the interactions; the second-based start and end of the interactions; and the emotional charge of the interactions. Interpersonal interactions in my data cover both verbal and non-verbal (e.g. nodding to answer a question) communication of team members.

The emotional charge of interactions refers to *how* the given interpersonal action relates to the previous action. I categorized the emotional charge of the interactions to the extent whether they are positive, negative (together

	start t	end t	send	rec	charge	sec	min	dur(s)
113	10:47:04	10:47:05	B	L	poz	121	3	2
114	10:47:04	10:47:05	B	V	poz	121	3	2
115	10:47:04	10:47:05	B	S	poz	121	3	2
116	10:47:04	10:47:05	B	N	poz	121	3	2
117	10:47:05	10:47:07	V	B	poz	122	3	3
118	10:47:07	10:47:07	L	N	neut	124	3	1
119	10:47:07	10:47:11	S	N	neut	124	3	5
120	10:47:07	10:47:11	S	L	neut	124	3	5
121	10:47:11	10:47:16	L	N	neut	128	3	6
122	10:47:12	10:47:13	N	S	neut	129	3	2
123	10:47:14	10:47:19	S	L	neut	131	3	6
124	10:47:19	10:47:21	L	N	neut	136	3	3
125	10:47:20	10:47:22	S	L	neut	137	3	3
126	10:47:20	10:47:22	S	N	neut	137	3	3
127	10:47:22	10:47:24	S	L	neut	139	3	3
128	10:47:23	10:47:29	L	N	neut	140	3	7
129	10:47:29	10:47:31	N	L	neut	146	3	3
130	10:47:31	10:47:34	L	N	neut	148	3	4

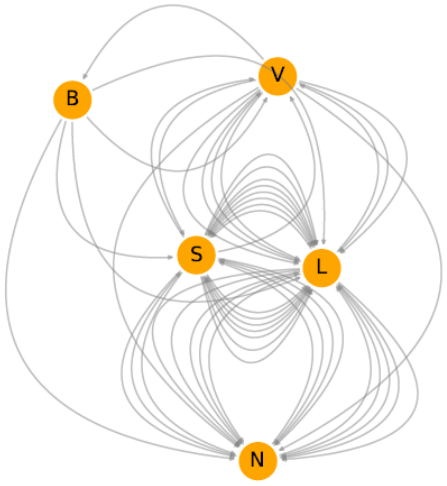


Figure 2.2. *Example of an edgelist extracted from video records and its network counterpart.*

referred to as signed) or neutral. Neutral interactions are directed to the problem-solving activity itself, and they are not loaded emotionally. Thus, regardless whether it expresses agreement or disagreement, communication is neutral until it is task-related. As soon as it becomes emotional, for example, the sender of the communication tie raises their voice, it becomes signed, negative in this case.

Signed interactions are those either related to social relationships or although task-directed, they are also emotionally loaded. For example, praising and the encouragement of one another are considered positive interactions as they are directed to social relations, and they are assumed to be constructive to these relations, thus influencing team spirit positively. On the other hand, negative edges are those that influence team spirit adversely such as creating tension, provocation or disparagement. Moreover, if a task-related remark is emotionally loaded, for example the interaction is directed to the task, but the sender of the communication tie is yelling, it is identified as a negative tie. This classification aims to mimic the well-established distinction between task-related and relationship-related social behavior [51][52][53].

In sum, neutral interactions cover emotionally indifferent, task-focused communication; while signed interactions are those influencing team spirit positively/negatively, as they are constructive/disruptive to social relations in

the context of a collaboration-demanding, pressed environment.

Another type of categorization - although not a separate parameter, - covers the distinction between dyadic communication ties and those addressed to more than one receiver, that is, group interactions. As video records are coded into an edgelist, group interactions are present in as many rows as many team members the communication was addressed to. In these rows, however, the start time, end time and the sender of the communication tie is identical.
Unfortunately, the above-detailed data can only be extracted manually. Therefore, I recruited and trained transcriber assistants for the transcription of the video records. In the transcribers' recruitment process, all candidates received the same five-minute trial video to code, along with a ten-page guide (see Appendices 6.0.4). In this guide, I described all details about how to extract the needed parameters properly. The trial video was crucial for two reasons: First, I gained information about the applicant's sense of precision and accuracy; and I could measure the inter-coder reliability by comparing the transcriptions of different candidates. The Krippendorff alpha test [54, 55] yielded acceptable values; they ranged from 0.67 (for emotional charge) to 0.78 (for sender). Second, the trial video was part of the selection process that enabled me to employ the most committed candidates. In the later stages of video transcription, I randomly checked the work of the transcribers, comparing the transcripts they made with the video recordings. Moreover, I meet them regularly to actively supervise their work and provide them with further training.

In the following, I present how I used this unique data to advance our understanding of the phenomena of human interactions and collaborative problem-solving. In particular, first, I investigate the role of team familiarity and cross-subgroup communication in the faultline diversity-team success relationship. Next, I explore how social interactions evolve during task performance, and how they relate to different team characteristics and success.

CHAPTER 3

FAULTLINE DIVERSITY, FAMILIARITY AND COMMUNICATION

3.1 Prelude

Most organizations fundamentally rely on teams [56, 57, 58, 15]. Teamwork allows groups to exceed individual task-solving capabilities, potentially increasing effectiveness and productivity. Diversity is considered a key characteristic of teams in organizations [12, 59], as it is a factor that can be easily manipulated to compose efficient teams, or to evaluate the fit of prospective team members [60]. As such, diversity is a central theme in organization research.

Research on the effects of diversity on team outcomes has yielded inconsistent findings. Surpassing the either-or outcome approach, where diversity has either a positive (e.g. [61, 62]) or a negative (e.g. [63, 59]) influence on team performance, more recently the focus has been shifted to more sophisticated models aiming to identify conditions under which diversity can be associated with certain outcomes [64]. A promising endeavor is marked by the faultline theory[1], which argues that instead of investigating the extent of group-level homogeneity of different individual attributes, diversity research should consider the joint, cumulative effect of different dimensions of diversity to better understand its negative influence on team outcomes [65]. Another notable development in this theme is the shift from linking certain types of diversities to specific team outcomes to identify moderators in the diversity-performance relationship [66]. While group faultlines have been investigated thoroughly in relation to various outcomes from team satisfaction [67] through creativity [68] to conflict and

[1]Hypothetical dividing lines that split the team into subgroups based on the alignment of different diversity dimensions, such as age, ethnicity, etc. [64]

trust [69], only a few studies have integrated faultlines theory with the investigation of moderators in the faultline-outcome relationship [70, 71, 72, 73]. Thus, it is still an understudied and yet fastly evolving area of diversity research [64].

In this chapter, we join and extend faultline research in the following ways. First, addressing the question of under which conditions diversity has a negative effect, and how faultline strength can be moderated on team performance, we test the moderating roles of two variables: the level of team familiarity and the ratio of cross-subgroup communication. Team familiarity has been found to have a direct positive impact on performance [74, 75, 76], however, its moderating effect in the diversity-performance relationship could only be detected in some cases depending on the nature of the observed diversity and task changes [77]. Here, we investigate the moderating role of team familiarity specifically, in the relationship of faultlines and performance. To quantify team familiarity, we measure the strength of intra-group relationships by looking at the duration of pre-existing social ties [78, 79, 80]. The concept of team familiarity contains the notions of trust, shared understanding, the knowledge of where expertise lies in the team, cohesion, and the sense of unity (e.g., [72, 81]. Thus, and in line with past research [74, 75, 76] we expect familiarity to have a substantial effect on performance, and to counteract the negative impact of faultline strength on team success.
The effect of communication has been investigated in relation to group faultlines [71, 82], and, as noted above, the moderating effect of faultlines has also been analyzed in some depth (e.g., [70, 73]. At the same time, cross-subgroup communication as a potential moderator in the diversity-performance relationship has not been explored. In this chapter, we address this matter. We expect cross-subgroup communication, as a proxy for the level of team-wide collaboration, to have a positive influence on team success. Moreover, we assume that the negative influence of faultline strength can be mitigated by interactions that cross the boundaries of those faultlines. In addition, in the faultline-performance relationship, we operationalize communication, for the first time to the best of our knowledge, as real-time interactions during the problem solving activity of small teams, as opposed to past research that almost exclusively rely on retrospective self-reported data. We employ survival models (also known as event-history models) that allows us to explicitly account for the time-varying nature of collaborative interactions during problem solving.

Our second contribution is connected to the collection of the data on real-time interpersonal interactions. In this regard, we introduce escape rooms as non-interventional social laboratories. Using this quasi-experimental setting, we collect data on the problem-solving activity of 40 project teams. As opposed to traditional laboratory experiments (e.g., [83, 26, 38]) that typically deal with

ad hoc, zero-history groups, these teams are intact, non-manipulated groups that are not torn out of their embedded social context. By analyzing real-time, collaborative interactions in escape rooms, we also overcome a shortcoming of most of the field studies that rely on self-reported data. Information on what happened in the team previously collected by retrospective questionnaires is likely to include biases such as highlighting the most salient event (e.g., interaction with a colleague) at the expense of all interactions [26, 38]. With our unique research design, we also aim to contribute to the increasingly rare studies of interacting groups that perform a practical task as opposed to just thinking about something. In addition to the above-mentioned advantages of this setting, escape rooms also present a cost-efficient research environment for capturing high-frequency real-time social interactions, a relatively sparse datatype due to the time-, effort-, and resource-intensive nature of collecting such information. Exploiting this more objective and novel source of data on real-time group interactions also allows us to disentangle the structure of intra-group relationships (the level of team familiarity in this case) from those of real-time communication ties – two distinct concepts that are often used interchangeably under the notion of social interactions [29, 2, 28]. Intra-group relationships are static and long-standing constructs tying members to one another [84], as opposed to actual interactions that are more dynamic and adaptive structures presenting a social means for effective task performance [33]. In a theoretical sense, we adopt both the categorization-elaboration model (CEM, [66]) and the network perspective to understand group processes through which diversity influences performance negatively.
We believe that both team familiarity and cross-subgroup interactions are crucial to understand teams as complex systems, as both variables account for core processes constituting complexity that makes teams more than simply the collection of individuals. However, these relations have clear differences in their very nature that we also aim to consider in this work.

Finally, our work answers to the calls for constructing and using more behavioral theory when studying operations [85, 86], and to pay attention to structural configurations of team processes (e.g.: subgrouping) [87].

3.2 Theoretical foundation

3.2.1 Diversity

Current trends in organizations show a diminishing level of verticality: institutions are flattening, and the decision-making processes are becoming less unidirectional. These developments have been fueling work systemized in team-

based structures, and more specifically, the rise of project teams [88, 89]. Agility and flexibility are increasingly valued characteristics of organizations in general, which strive to keep up with their rapidly changing environments, and, in the case of firms, in particular, to meet the varying demands of the market. Project teams are the group-level manifestations of organizational agility and thus, institutions rely more and more on them. The term 'project team' refers to 'an organized, task-focused group' [84] (pp.352.), while it also denotes the non-routine and unpredictable nature of the task to be accomplished by the team [90]. In addition to organizations' growing reliance on teamwork, demographic changes and growing specialization in the workforce coupled with the increasing mobility of employees between workplaces [91] make organizations more and more heterogeneous [12, 92]. Thus, team diversity is a recurrent factor in organizational life and performance.

Team diversity refers to the extent to which team members differ from one another [12]. Studies on team diversity typically focus on heterogeneity in gender, age, tenure, functional background, and ethnicity [93, 94, 95]. One of the central questions of team researchers is how team diversity as a compositional factor influences performance [96]. Scientific literature provides conflicting answers on this matter. One stream of research rooted in social psychology considers diversity detrimental to performance [63], as it involves an automatic classification of people into categories by which the notion of "we" and "they" evolves immediately [97]. Thereby, the social categorization process "provides a cognitive foundation for intergroup conflict." [84] pp.422. The pervasiveness of intergroup bias is best demonstrated by the research of Henri Tajfel et al. [98] on what they called a minimal (inter)group situation. They found that even in ad hoc groups without prior group history, members, who have not known each other, expressed partiality toward members of their group based on superficial similarities. Adopting this perspective, several studies claimed that social categorization can be disruptive to the team's cohesion by negatively influencing intra-group relationships, which can cause process loss and mismanagement [99, 100, 101, 66].
In contrast to this mechanism, the information processing/cognitive resource theory born in the field of management sees diversity as beneficial to team performance. This perspective considers diversity as an informational resource, based on the assumption that differences between people entail differences in knowledge, experiences, skills, etc. Thus, a greater amount of team diversity may be associated with a greater variety of cognitive resources members can bring to the team. Heterogeneous membership of teams is assumed to promote creativity and innovation by utilizing the profusion of aspects. Rich and diverse inputs of team members also hypothesized to improve performance by

eliminating the trap of groupthink [102]. Consistent with this approach, diversity studies reported about the advantageous nature of heterogeneity on team performance [103, 61, 104]. As plenty of evidence supports both outcomes of diversity [66], research has been inconclusive in identifying unique and direct negative and positive effects of different diversity dimensions on performance (e.g. [94, 105].

3.2.2 Faultlines and subgroup formation

In the quest to establish coherence in diversity research, a stream of studies turned to investigate the joint effects of multiple types of diversity on team outcomes [64, 106]. In contrast to the traditional approach focusing on single attributes (such as gender or age), this stream of research examines faultlines, that is, "hypothetical dividing lines that may split a group into subgroups based on one or more attributes" [65] pp.328). Therefore, faultline research operates with the cumulative effects of attributes on team outcomes. The theoretical underpinnings of group faultlines lay in the social categorization theory, which claims a strong homophilic tendency of individuals. Individuals' alignment with similar others based on demographic traits can result in subgroup formation that disrupts information exchange and performance 3.1.
Further explaining the mechanisms of diversity faultlines, the categorization-elaboration model (CEM) [97] draws attention to the notion of comparative fit. Comparative fit refers to the extent to which the categorization of differences manifests in high intra-subgroup similarity coupled with high inter-subgroup dissimilarity. For example, when younger members of the team also tend to be women, we can say that these positions on the age and gender dimensions of diversity are correlated. The more these correlations are, the higher the comparative fit of categorization in these dimensions (age, gender – younger women, older men), and the more likely that it will induce strong faultlines that divide the team into subgroups. In other words, the presence of a strong faultline in the team accounts for subgroups that are induced along different dimensions of diversity, in a way that members of these subgroups are highly similar to each other while highly different from members of the other subgroups. Strong faultlines explained by CEM were found to disrupt group processes and hinder success (e.g. [107, 108, 95, 109].

More recently, the focus has been also shifted to detecting moderators of the diversity-performance relationship to understand under which conditions diversity implies certain outcomes [66]. To this end, although framing differently, both CEM and the network perspective [28, 60, 87] suggest that influence of diversity can be understood via the examination of team processes. CEM lays emphasis on understanding the process of social categorization (associated with

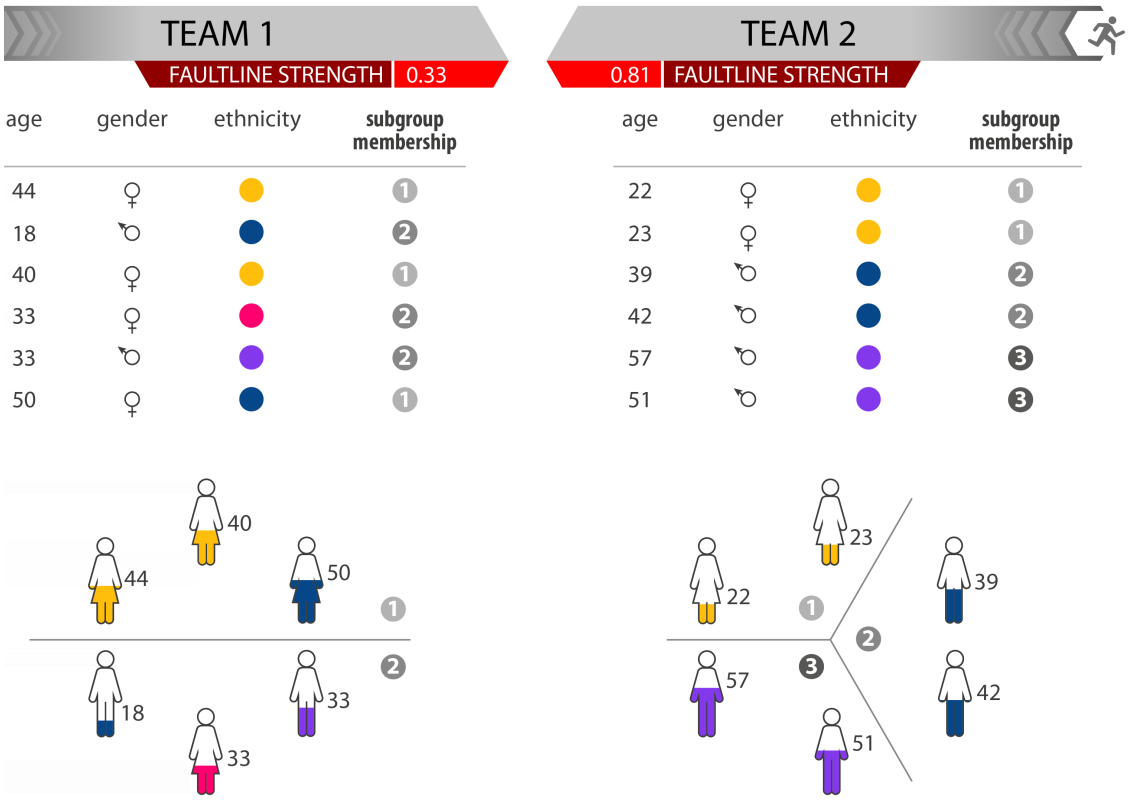


Figure 3.1. ***Visual presentation of faultline strength.*** *As we can see, Team 1 and Team 2 have a similar level of diversity considering diversity dimensions of age, gender and ethnicity separately. However, when observing these dimensions in a cumulative manner, Team 2 has a significantly higher* 0.81 *faultline strength. The subgroup formation provides an intuitive understanding on what this value expresses. It is conspicuous that the hypothetical dividing lines splitting Team 2 into subgroups results in undoubtedly more homogeneous subgroups than in the case of Team 1. Thus, the more homogeneous intra-subgroups and more heterogeneous inter-subgroups are, the higher the faultline strength is. Therefore, it is assumed that Team 2 is more likely to experience process losses and erratic communication patterns than Team 1.*

the negative effect of diversity) in order to both prevent the negative effects and provide the preconditions for the positive impact of diversity [97]. On the other hand, the network perspective claims that different forms of social capital (operationalized as ties of social interactions) regulate the strength of association between team diversity and productivity [28]. The common, underlying con-

cept of these theoretical streams is that those seeing diversity as detrimental to performance expect this relation because diversity can erode groups' social cohesion and induce coordination problems, process losses, and insufficient information exchanged via the process of social categorization. Instead, the information processing perspective draws attention to the positive influences of diverse knowledge and ideas within groups. Still, members' diverse cognitive resources present a potential rather than a promise for successful problem-solving as long as the group-level synthesis of individual assets is overlooked. Our research has been informed by these perspectives, which we adopt and rely on. We build on the comprehensive model of CEM that is well-supported by empirical evidence (e.g. [107, 108, 110], and integrate it with the faultline approach. Although we recognize the added value of faultline theory over that of diversity, we follow the recommendation of Bezrukova et al. [111, 112] and Lau and Murnighan [71] to include diversity effects into the analysis with the aim to isolate the unique effects of faultlines connected to a specific member alignment.

3.2.3 Team familiarity and the strength of intra-group bonds

Faultlines induced by the alignment of team members' demographic traits increase the likelihood of subgroup dynamics, and thus, threatens social integration [65]. In other words, teams with strong faultlines are more exposed to the process of social categorization. However, the perceived importance of dissimilarities can be mitigated by the time group members spend together [113]. Indeed, team-building adventure is a popular method by which organizations aim to create a sense of unity in their teams. Factors that promote familiarity such as proximity or (deep-level) similarity are often discovered through shared experiences and tasks to be solved, which enhances strong relationships among team members and thereby makes the team more cohesive [84].
The unity that these practices aim to create accounts for the added value of teams that makes them adaptive and efficient task-performing systems surpassing individual capabilities. This advantage of teams, specifically that teams are more than merely the collection of their members' traits and capabilities is rooted in their complexity [35]. Therefore, we believe that to understand teams in their complexity, it is imperative that we investigate their structure – the system of intra-group relations that outlines the quality of groupness. In this respect, and considering teams as organized systems of relationships, we focus on the strength of bonds connecting members to and in the group. Based on prior research, we assess the strength of relationships by the length of acquaintanceship between members. The duration of relationships is one of the most reliable indicators and an essential dimension of relational strength [78, 79, 80]. Long-term relationships indicate a great amount of knowledge about each other

that implies predictability and thereby, trust between members [114, 115, 116]. Team familiarity that is prior interpersonal knowledge about other members [117, 15] can reduce intra-group biases by changing members' perception of group boundaries, and thereby help members to perceive themselves as part of a unit with a shared identity [118]. Thus, team familiarity enhances cohesion by decreasing the perceived importance of intra-group faultlines [72]. In addition, Huckman and Staats (2010) found that team familiarity facilitates the management of the differences in prior experience of group members. Combining these findings, we assume, that in teams having a strong faultline and thus, a presumably more pronounced process of social categorization, the elaboration of information is less smooth compared to teams with weaker faultlines. However, the knowledge familiarity presents can foster the team's collective processing of individual task-related ideas. Therefore, we consider team familiarity a mechanism that may well improve coordination and thereby, enhance team performance.

Regarding familiarity's main effect on performance, a body of research [74, 75, 76] observed that team familiarity also has a positive observed that team familiarity has a positive impact on team success, as it provides members with knowledge of who knows what that helps them to locate expertise and operate in a common context. The existence of this knowledge and shared understanding is particularly important in project teams that face an innovative, complex, and non-routine task as the lack of predictability of the task makes members highly interdependent and thereby resort to team familiarity [75]. In other words, strong relationships hereby defined as long-lasting social ties make members less dependent on their own knowledge and decrease uncertainties induced by social dilemmas and interdependencies. Team familiarity promotes a shared understanding of the work, facilitates the avoidance of interpersonal risks, and provides members with collective responsibility for the task. It also implies a basic psychological safety in which members may incline to voice even their half-finished thoughts as they do not risk their reputation or being judged negatively [119]. Navigating ambiguous and creative tasks, it is essential for project teams to feel safe enough to take risks and communicate honestly.

Based on the above reasoning, we assume that a strong web of (long-term acquaintanceship) relations accounting for a high level of team familiarity positively influences problem-solving. We also expect that mutually accustomed members perceive subgroup boundaries as less important as they had time to work through their initial differences potentially induced by the process of social categorization. The level of team familiarity implies a shared understanding, trust, and the feeling of unity, and thus, it is assumed to moderate the potentially negative impact of faultlines on task performance. We formulate

our related hypotheses in the following way:

H1a: *Familiarity has a positive impact on team success.*

H1b: *Familiarity moderates the negative effect of faultline strength and diversity on team success.*

3.2.4 The role of team-wide interactions

As outlined by CEM, subgroup formation induced by differences endangers team processes by disrupting information exchange and integration. Thus, strong faultlines, that is, when attributes of members are correlated (e.g., the subgroup of women also tend to be young and highly educated) have a negative effect on the quality of collaboration, while, without engendered subgroup formation, information sharing is assumed to be smooth and effective, and thereby supporting performance [66, 95].
Information sharing is at the heart of the collaborative problem-solving processes. Indeed, Earley and Mosakowski [120] found that team communication reflecting the extent of information elaboration mediates the relationship between diversity and performance. Faultline researchers have also begun investigating moderators of the faultline-performance relationship. According to Thacher and Patel (2012) [64], only a few moderators such as superordinate identity or faultline distance have been investigated to a certain extent [72]).
Although communication has been addressed in relation to faultlines and group learning [71, 82], it has not been examined as a real process moderator in the faultline-performance relationship. Instead, the moderating role of faultline strengths has been brought to the fore and showed to have a negative impact in some aspects of the communication and performance nexus [71, 82].

Here, we contribute to the evolving understanding of what intensifies or attenuates the relationship between faultlines and outcomes by focusing on the role of actual communication.
The direct role of communication in team performance has been extensively examined by researchers following a functional and network perspective to study small groups [121]. From a network perspective, a group can be considered as a network, where ties are interactions between members. Field studies concluded that actively communicating groups outperform groups with fewer ties [27, 28]. The exchange and integration of information is a particularly critical process to performance in the case of project teams dealing with non-routine tasks demanding close collaboration (e.g. [122, 123]. Due to the increased interdependence and the collaboration-demanding nature of the task, group members are compelled to engage in intense information exchange and discussion [118]. Un-

der these circumstances, and sharing a common goal, even members who are reluctant to cooperate due to perceived intra-group differences become inclined to interact to achieve success [124, 125, 126, 95]. Consequently, research suggests that intense communication might be enough to overcome difficulties related to the alignment of team members' demographic attributes (such as conflicts, coordination problems, the lack of shared identity, etc.) by attenuating the effect of diversity on team success. Applying these finding to the faultline-performance nexus invoking CEM, we believe, that problem solving communication presents, strictly speaking, one of the most salient group processes through which diversity can affect team outcomes. Actual communication is the very social means team members can exploit to elaborate ideas and integrate fragmented knowledge of individual members.

At the same time, much research on collective action assumes that all team members communicate simultaneously with each other. However, even in the case of strategizing tasks when people sit around a table and discuss a strategy, this is often not true. Team members are typically linked to specific others via network ties [81]. Moreover, the pace and distribution of these interactions are often highly asymmetric across time of task performance and group members.

Addressing the inherently structural nature of teamwork is supported by the network approach and not without precedent. For example, LePine [127, 128] investigated the adaptation of role structure to task-focused interactions as a response to unforeseen changes making habitual interactions problematic. Similarly, other research studied communication patterns and roles in decision-making processes to understand structural adaptation (e.g. [129, 130]. Resonating to the notion of the structure of communication, the social categorization and social identity approach claims that group members prefer to interact with similar others. In strong faultline groups where members can be split into highly homogenous subgroups, team members will favor to communicate with members of their own subgroup, "us", instead of "them", the subgroup including members perceived as different. In this case, the elaboration of information does not manifest on the level of the team. In other words, the team as a unit fails to exploit the potential in its complexity, and as such, it cannot function efficiently. Thus, similarly to Lau and Murnighan [71], we focus on communication between the faultline-induced subgroups. At the same time, instead of focusing on faultlines as moderators, we analyze the moderating role of cross-subgroup interactions in the faultline-performance relationship. We believe that complexity is the strength of teams making them more than the collection of individuals, and this complexity is encoded in team processes, such as communication. When teams manage to transcend subgroup boundaries and thereby induce an active, team-wide communication, we expect this process to mitigate

the negative faultline effect on performance.

Most empirical studies attempting to integrate communication with the notions of diversity and performance, obtain communication data from self-reports, where participants are asked to evaluate their group's actions, interactions, and outcomes [71]. This data is likely to contain biases highlighting the most salient or recent event at the expense of all events (e.g., communication with a colleague) [131]. Furthermore, interactions in this sense are the proxy of relational strength [29, 29, 28], rather than ties covering real interactions between team members – two types of relations we address in this work while considering the differences in their very nature (e.g.: temporality and dynamics). Consistent with Carton and Cummings's [132] and Crawford and LePine's [87] calls for researchers to consider the structure of taskwork, we explicitly take into account the structure of collaborative interactions assessing whether they occur within or between the faultline induced subgroups. Moreover, in our work we do not rely on retrospective questionnaire data of team members memory and necessarily biased perceptions but analyze real-time communication to understand the collaborative dynamics of team activity.
Based on the above-mentioned research detailing the importance of communication both as a predictor and moderator, we expect that communication may counteract the potential negative effect of faultlines on performance when collaborative interactions tend to transcend subgroup boundaries. Otherwise, the team-wide information exchange and group-level integration of task-related knowledge do not occur, and faultlines affect performance negatively via social categorization. Therefore, the non-routine nature and high collaboration demand of the task might be not properly met, and thus, both faultlines and within-subgroup communication can directly express their negative effect on performance. Besides regulating the strength of association between faultlines and performance, we expect a high level of cross-subgroup communication to have a beneficial impact on team success. We summarize our related hypotheses as follows:

H2a: *Cross-subgroup communication has a positive impact on team success.*

H2b: *Cross-subgroup communication moderates the negative effect of faultline strength and diversity on team success.*

3.3 Concepts and Measures

Team success. The very essence of any teamwork is to cooperatively implement a task. Therefore, we were interested in measuring successful problem-solving.

We operationalized success as follows: teams that did not manage to escape within an hour were considered failed groups, while those who did escape within the available time frame were labeled as successful teams.

Diversity. Although we recognize and build on the added value of faultline theory over that of diversity, we follow the recommendation of Bezrukova et al. [111, 112] and Lau and Murnighan [71] to include diversity effects into the analysis (as separate control variables) to isolate and compare the unique effects of faultlines with the traditional diversity measures. Concerning team diversity, we primarily focused on the most frequently used social-category differences such as age, gender and the highest level of education of team members. For the nominal predictor, gender, notation one accounted for men, while two for women. We measured the highest level of education on a one to four ordinal level scale, where one denoted elementary, while four indicated higher education (college/university degree). As both of these variables are categorical levels, parallel with other diversity research [59, 95], we used Blau's [133] index to capture diversity (as variety) in gender and education background. The continuous variable age was measured by years, and its diversity measure was operationalized by the standard deviation [59, 95].

Faultline. Among the several methods to calculate faultline strength, we opted for the average silhouette width (ASW) relying on a cluster-analytic process [134]. ASW measures the same construct as the most widely used faultline measure, Fau [64]. Additionally, ASW allows to determine more than two subgroups. According to Meyer and Glenz [134], ASW is the average of team members' individual silhouette width, representing how well a group member i fits into cluster A compared to cluster B. Individual silhouette width is expressed by the following formula

$$s(i) = \frac{b_i - a_i}{max(a_i, b_i)} \tag{3.1}$$

where a_i is the average dissimilarity of i to all members of cluster A,
b_i is the average dissimilarity of i and all members of cluster B.
Dissimilarities are measured by the Euclidean distance between individuals.

Based on a comprehensive empirical comparison by Meyer and Glenz [134], and Meyer et al. [135], this is the most robust method for detecting faultlines.

We calculated ASW across the three most frequently considered dimensions of diversity, that is, age, gender, and education.

The level of team familiarity - group structure. Team familiarity as a structural factor is requisite for a favorable team environment and is hypothesized *(H1)* to moderate the effect of faultline strength and diversity on team success. To assess the level of group familiarity, we decided to examine the strength of pre-existing social ties [101], and in particular, the length of acquaintanceship [78, 79, 80]. We used the questionnaire to gather this relational and structural data by asking each member how long (s)he has known each team member measured in years. The answers outlined the strength of intra-group relationships. The level of familiarity was measured by the mean of the years team members have known each other. Teams, having a low average value of the length of acquaintanceship possess a lower level of group familiarity, while a high level of familiarity is present in groups with high values of the average length of acquaintanceship.

Interpersonal interactions - the ratio of cross-subgroup communication. We also incorporated teams' real-time communication. As we have mentioned, each real-time communication was coded from the video records. Here, we focus on the interpersonal interactions among the subgroups potentially created by the faultline. Therefore, first, we calculated the cumulative sum of all communication for each minute of the game. Then we determined the number of communication ties that occurred between actors of different subgroups defined by the faultlines. Then, we calculated the ratio of cross-subgroup communication, that is the division of the cumulative sum of cross-subgroup interactions by all communication ties for each minute. This measure shows how (relatively) frequently group members break the faultlines in their communication, by jumping through the homogeneous boundaries.

Type of the escape room. As our data come from two escape rooms, we also included the variable of the room identifier into the analysis to make sure that the results are not contingent upon which room the teams performed the task. The identifiers of rooms Sherlock and Godfather were zero and one, respectively.

Table 3.1 provides descriptive statistics about the dependent variable success, the predictors, and the control variable. Our data is not perfectly balanced in the sense that we have more failed groups than successful ones, and the majority of the teams played in the room Sherlock. Successful teams needed 57 minutes on average with small variations to escape either of the rooms. We can

also see that the average Blau index of gender within a team is 0.37 on average, which suggests that teams in our sample are rather diverse. The average age of the participants in our sample is 31 years old, and the average standard deviation of age in a team is 4.5 years.

	mean	**std**	**min**	**max**
time of task performance (min)	57.7	3.884	47	60
success	0.35	0.483	0.0	1.0
room	0.35	0.483	0.0	1.0
gender (Blau's index)	0.366	0.189	0.0	0.5
age (avg)	30.903	8.743	18	50.5
age (std)	4.465	4.036	0.0	16.19
education (Blau's index)	0.315	0.246	0.0	0.64
ASW	0.508	0.235	0.0	0.966
acquaintanceship (avg)	7.059	6.905	0.38	35.08
ratio of cross- subgroup interactions	0.589	0.161	0.0	0.866

Table 3.1. *Descriptive statistics (team-level)*

The average Blau index of the highest level of education within teams is similar to that of gender, namely 0.315. This value suggests rather diverse teams in terms of education background. On average, team members have known each other well (approximately 7 years), however, the variation of average acquaintanceship across teams is relatively high (6.9). The ASW values in our sample range from zero to one, where zero means no hypothetical dividing lines, while one represents strong faultlines [106]. The average ratio of cross-subgroup communication is 0.59, showing that on average around 60% of all collaborative communication occurs between the faultline-induced subgroup, with a variation of 0.16. Both ASW values and the ratio of cross-subgroup interactions are normally distributed.

3.4 Survival models

We applied multivariate models in the analysis. We employed discrete-time survival (also called event-history) models [136] with both time-fixed and time-varying variables, to investigate the direct and moderating effects of team familiarity and cross-subgroup communication. We used a time-team level data structure, where a team is present in the dataset in as many rows as many minutes they spent with problem-solving. The binary variable, success has the

value of zero in each minute the group performed the game, and one in the minute the group succeeded by leaving the room. Time-fixed variables, like diversity measures, are constant in each minute of the same group, while time-varying variables, like the cumulative ratio of cross-subgroup interaction, are changing over time within the group. By employing survival models, we could incorporate the dimension of time into our investigation, by taking into account the time-varying nature of the process-related variables.
In the first model group (Model 1,2,3,4) we gradually included predictors into the analysis to investigate their direct effect on team success. Model 1 is the baseline model, which shows how much the probability of exiting the room is raised by the addition of one minute for task accomplishment without considering any team-related variables. Model 2 additionally includes the time-fixed, compositional predictors, namely diversity and faultline measures, and the control variable of the room. In Model 3, we extend Model 2 with the relational variable of familiarity, and in Model 4 with the process-and-structure-related predictor of the ratio of cross-subgroup interactions to test H1a and H2a. Then, to Model4 we add interaction terms to examine the interplay of familiarity and cross-subgroup communication with diversity and faultline measures. We tested the potential moderation role of familiarity and cross-subgroup communication on the effect of diversity and faultline measures on success *(H1b, H2b)* by these models.

First, we created the baseline model (Model 1) to examine the relationship between the minute-based problem-solving time and success. Then, in Model 2 we included faultline strength and diversity measures. The latter also serve as controls for the faultline. We also added the type of escape room as a control variable. Model 3 contains the relational predictor (familiarity) too, while the most comprehensive model, Model 4, involves the ratio of cross-subgroup communication as well. We summarized the results in Table 3.2. All four models are significant at a 0.001 significance level. With the increasing complexity of the models, from Model 1 to Model 4, their predictive power, according to the Tjur R2, also increases within the range of 0.021-0.096.

Predictors	**Model 1** *Log-Odds*	**Model 2** *Log-Odds*	**Model 3** *Log-Odds*	**Model 4** *Log-Odds*
(Intercept)	-12.46^{***}	-8.72^{*}	-9.78^{**}	-9.77^{**}
	(2.21)	(3.59)	(3.58)	(3.59)
time (min)	0.16^{***}	0.21^{***}	0.23^{***}	0.23^{***}
	(0.04)	(0.05)	(0.05)	(0.05)
average age		-0.10	0.02	0.02
		(0.07)	(0.08)	(0.08)
age diversity		0.01	-0.33	-0.33
		(0.14)	(0.22)	(0.22)
education diversity		-4.68^{**}	-4.40^{**}	-4.35^{*}
		(1.75)	(1.63)	(1.89)
gender diversity		-2.70	-4.52^{*}	-4.48^{*}
		(1.56)	(1.85)	(2.02)
faultline strength		-2.41	-3.20	-3.17
		(1.66)	(1.87)	(1.99)
room		0.39	0.10	0.11
		(0.62)	(0.63)	(0.67)
average acquaintanceship			-0.26^{*}	-0.27^{*}
			(0.11)	(0.11)
ratio of cross-subgroup interactions				-0.14
				(2.77)
Observations	2308	2308	2308	2308
R^2 Tjur	0.021	0.081	0.096	0.096
		$^{*}p < 0.05$	$^{**}p < 0.01$	$^{***}p < 0.001$

Table 3.2. *Discrete-time survival models on exiting the room (success).*

As the baseline model suggests, providing more time for task accomplishment raises the probability of success. The effect of time on exiting the room increases when we include compositional variables (Model 2), and increases further when incorporating the relational and communication characteristics of the teams (Model 3 and Model 4). In Model 2, we tested the impact of different dimensions of diversity and faultline strength on success. Keeping all the other predictors constant, Model 2 only showed a significant (and negative) effect in the case of education diversity on team performance.

In Model 3, we included the average acquaintanceship, our measure for group familiarity. Surprisingly, the level of team familiarity has a negative impact on success. Moreover, the average length of acquaintanceship has a suppression effect, as its inclusion into the model induces the predictive power of

gender diversity. Thus, in those groups where the level of team familiarity is the same, teams with lower gender and education diversity are more prone to succeed. Similarly, in teams with the same level of gender and education diversity, the less familiar ones tend to be efficient problem-solvers. Therefore, we reject H1a, as Model 3 suggests that familiarity does not have a positive, but a negative effect on team success.

In Model 4, we tested the direct effect of cross-subgroup interactions on team success (H2a), which turned out to be non-significant. As Model 4 is identical to Model 3 in terms of the signs and significance of the coefficients, and as the direct effect of cross-subgroup interaction was not significant, we also reject H2a, so the minute-based cumulative ratio of cross-subgroup communication does not have a direct influence on exiting the room. In sum, faultline strength is not detected as a significant predictor of team success. However, gender and education diversity as well as the level of team familiarity proved to have a meaningful, negative influence on team performance. In order to understand this intriguing and somewhat unexpected message and testing the potential moderation effects of both relational and communication features on diversity and faultline measures (H1b, H2b), we added the interactions of these features to the last, most complex survival model.

First, we explored if the interactions between team familiarity and team composition variables (faultline and diversity indicators) have significant effects on success, and so whether familiarity modifies the impacts of diversity and faultline strength on success (H1b). The results are summarized in Table A1 (in Appendices 6). We have not found any significant interaction between familiarity and diversity or faultline measures. We found that the negative effect of gender and education diversity on performance is present regardless of how long team members have known each other on average. Moreover, age diversity and faultline strength are still irrelevant to success in teams both with high and low average familiarity. Therefore, in none of these models does the level of familiarity leverage the effect of diversity or faultline strength on success. As a consequence, our expectation on the moderation effect of familiarity in the relationship of diversity or faultline with success (H1b) is not confirmed.

We built three other models with interaction terms between diversity and faultline strength, and cross-subgroup communication, to test whether communication that bridges faultline-created subgroups has a moderating role in the relationship of diversity and faultline with success. We summarized the results in Table 3.3.

Predictors	**Model 5** *Log-Odds*	**Model 6** *Log-Odds*	**Model 7** *Log-Odds*	**Model 8** *Log-Odds*
(Intercept)	-13.98^{***}	-10.03^{**}	-9.71^{**}	-5.84
	(4.12)	(3.72)	(3.63)	(4.19)
time (min)	0.27^{***}	0.23^{***}	0.23^{***}	0.27^{***}
	(0.06)	(0.06)	(0.05)	(0.07)
average age	0.06	0.02	0.02	-0.01
	(0.08)	(0.08)	(0.08)	(0.10)
age diversity	0.37	-0.33	-0.34	-0.36
	(0.25)	(0.23)	(0.23)	(0.25)
education diversity	-5.12^{*}	4.61^{*}	-2.14	-4.09^{*}
	(2.14)	(2.17)	(6.22)	(1.96)
gender diversity	-4.69^{*}	-3.03	-4.82^{*}	-6.44^{**}
	(1.98)	(6.11)	(2.30)	(2.44)
faultline strength	-3.67	-3.42	-3.67	-10.77^{*}
	(2.03)	(2.22)	(2.54)	(4.66)
room	-0.40	0.07	0.05	0.17
	(0.71)	(0.69)	(0.69)	(0.67)
average acquaintanceship	-0.26^{*}	-0.26^{*}	-0.27^{*}	-0.42^{**}
	(0.11)	(0.12)	(0.11)	(0.14)
ratio of cross-subgroup interactions	4.97	0.88	0.84	-6.35
	(3.69)	(4.98)	(3.96)	(4.34)
ratio of cross-subgroup interactions*age diversity	-1.69^{**}			
	(0.63)			
ratio of cross-subgroup interactions*education diversity		-2.82		
		(11.38)		
ratio of cross-subgroup interactions*gender diversity			-4.22	
			(11.59)	
ratio of cross-subgroup interactions*faultline strength				13.04^{*}
				(6.54)
Observations	2308	2308	2308	2308
R^2 Tjur	0.114	0.097	0.096	0.122
		$^{*}p < 0.05$	$^{**}p < 0.01$	$^{***}p < 0.001$

Table 3.3. *Discrete-time survival models on exiting the room - models with interactions.*

As can be seen in Table 3.3, the interactions of cross subgroup communication with education and gender diversity are not significant (Model 5 to Model 8). However, we detected significant interaction between cross-subgroup communication and age diversity, as well as cross-subgroup communication and faultline strength. Model 5 contains the interaction term between cross-subgroup communication and age diversity, suggesting a cross-over effect between these predictors on success. Indeed, none of the two predictors has a significant direct impact on success, unlike the interaction of the two which, to our surprise, is found to be negative. To reveal where this negative effect lies, we portrayed the average marginal effects on success for each combination of low/high age diversity with low/high cross-subgroup communication (Figure 3.2). The negative effect of age diversity is much stronger in homogeneous teams where members tend to initiate a small amount of communication across subgroups, compared to age-homogeneous teams with a high ratio of cross-subgroup communication. Similarly, although with smaller differences in the effect sizes, teams with high age diversity tend to perform better when they communicate within subgroups.

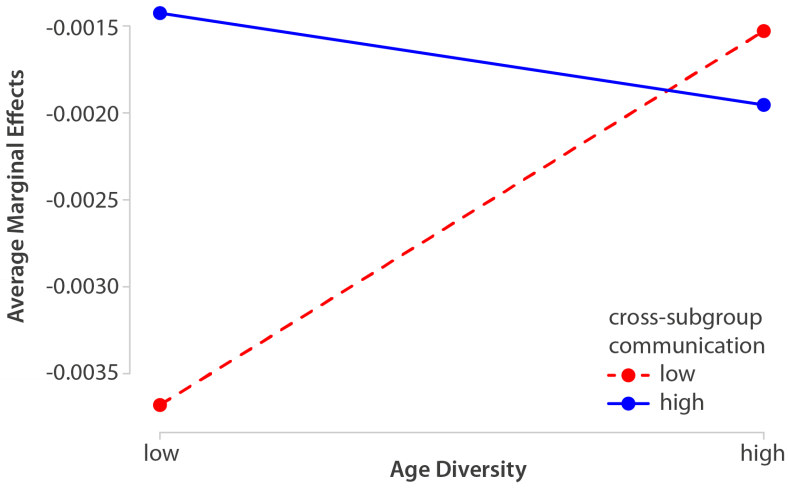


Figure 3.2. *Interaction plot on success by cross-subgroup interactions and age diversity.*

According to the interaction term, the negative effect of age diversity on success can be minimized in different ways: either if team members are close to each other in age (low age diversity) and communicate a lot across subgroups; or if the team is diverse in terms of age, and members communicate within rather than across subgroups.

Model 8, which includes the interaction of cross subgroup communication and faultline strength, also shows a significant interrelation but of different signs (Figure 3.3). In teams with weak diversity faultline, low cross subgroup communication is a better strategy to succeed than communicating a lot across subgroups. The differences are more pronounced when we observe teams with strong faultlines. In these groups, the negative effect of faultlines can be attenuated by a high ratio of communication across these faultline-induced subgroups.

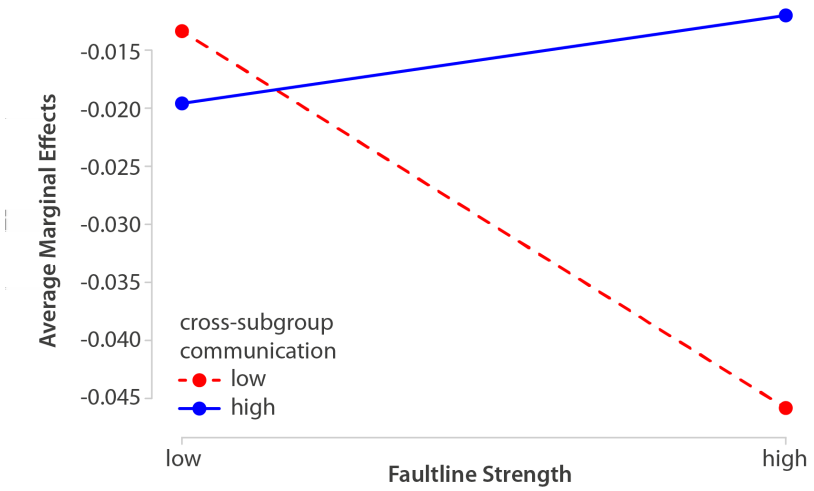


Figure 3.3. *Interaction plot on success by cross-subgroup interactions and faultline strength.*

The worst-case scenario for performance is when these teams with strongly correlated dimensions of diversity do not communicate enough across subgroups. In other words, faultline strength worsens the chances for project teams to succeed when the subgroups induced by the faultline do not communicate on an above-average frequency with each other. It also implies that when teams manage to elaborate ideas and integrate knowledge on the group level by crossing subgroup boundaries with communication, it can counterbalance the negative faultline effect. Thus, we found support for H2b, as the analysis confirmed that intense communication across the subgroups do moderate the negative effect of diversity on team success. Indeed, we identified cross-subgroup communication as a specifying factor in the relationship of faultline strength and age with performance and detected the circumstances under which the negative effects occur, and those situations where they do not. Taken together, the performed survival analyses shed light on several relationships between various

predictors and team success. Education diversity proves to have a negative influence on performance. Whereas, gender diversity shows the same kind of impact, only when the length of acquaintanceship is taken into account. Although cross-subgroup communication does not have a direct effect on success, it has a notable role in moderating other predictors' effects. Teams with low age diversity do better with more cross-subgroup communication. Most importantly, we found that the destructive effect of faultlines can be revealed only in the light of communication structure: diversity faultline is disruptive to success when the faultline-induced subgroups do not communicate on an above-average frequency, while in teams that do, the effect of faultlines is much weaker – close to being insignificant.

3.5 Discussion

This chapter joins and contributes to faultline research by investigating the direct and moderating effects of team familiarity and cross-subgroup communication in the diversity faultline-success relationship.

Building on the strength – but overcoming disadvantages – of both laboratory experiments and field studies, we introduce escape rooms as minimally biased social laboratories, free from the typical weaknesses of conventional experiments. Exploiting this innovative setting, we collected fine-grained data on the real-time communications of team members. In parallel, we explored the wider social embedding of individuals by questionnaire data that informed us about teams' composition and social structure.

Relying on the CEM and network perspective, we argued that a strong web of relationships might help to offset the importance of perceived intragroup biases induced by subgroup formation. We measured relational strength by the length of acquaintanceship to determine the level of team familiarity. In line with previous studies, we first hypothesized that team familiarity positively influences success (H1a), and it moderates the impact of faultline strength (H1b). Contrary to our expectations, survival model analyses indicated a negative influence of familiarity on performance. Moreover, the models revealed the suppression effect of familiarity, as its inclusion into the analysis activated the negative effect of gender diversity on team success.

A potential explanation for failing to confirm the familiarity-related hypotheses (H1a, b) lies in the operationalization of the concept. In this work, we considered prior interpersonal knowledge about other members [117, 15] as a proxy for team familiarity. This variable captures a general form of familiarity that cannot distinguish between different dimensions of it, such as prior work experience with the same crew [137, 138], or prior work experience with

the same teammates in similar prior tasks [139, 140, 74]. However, our analysis demonstrated a significant, but negative effect. Considering the non-routine nature of the task to be performed under time pressure in a dynamic environment, members must adapt to the varying demands of the task rapidly. In these circumstances, familiar team members may become stuck in routinized patterns of communication that is rooted in the strength of relationship. This may hinder them to efficiently align their collaborative behavior with the dynamically changing demands of the task [141, 142, 143]. Katz and Allen [144] showed such negative effect in an industrial setting, at teams who have worked together for a longer time. Similarly, familiar teams in our data have known each other for years, presuming well-established communication channels based on strong relationships. Moreover, the type of these relationships, (e.g.: friendship or coworkership) may also influence the direction of familiarity's effect on performance, as Harrison et al. [15] hypothesized an effect that is more likely to be negative in case of friends. Further research should address these dimensions of teams' social relations to clarify the effect of member familiarity on team success.

In line with the social categorization perspective, we detected negative influences of different dimensions of diversity. However, faultline strength proved to be insignificant in terms of performance. Nevertheless, incorporating into the analysis our process-related predictor, namely the ratio of cross-subgroup interactions, provided us with some new and intriguing insight. First, we identified a cross-over effect of our last composition factor, age, and the cross-subgroup interactions. The model suggested that age diversity can negatively affect success in two ways: either when age-homogeneous groups have a high amount of cross-subgroup communication or when age-heterogeneous teams tend to initiate within subgroup interactions. More importantly, we found support for the attenuating power of cross-subgroup communication in the faultline-outcome relationship (H2b). While the main effect of cross-subgroup communication remained insignificant, similarly to the previous interaction term, we detected a cross-over interaction between faultline strength and cross-subgroup communication. A high communication ratio across the faultline-induced subgroups attenuates the negative influence of strong diversity faultlines. However, when teams with solid faultlines do not surpass these dividing lines via problem-solving interactions, they tend to fail. Therefore, faultline strength can negatively affect team outcomes *when* it obstructs team level communication so that collaborative interactions typically occur within the faultline-induced subgroups. Naturally, when teams are characterized by non-existing or weak faultlines, a low amount of cross-subgroup communication is expedient. Yet, in the presence of strong subgroup formation, communication across these groups no-

tably increases the likelihood of success. In short, the effect of diversity faultline and cross-subgroup communication on success can only be understood in relation to each other. We show that cross-subgroup communication moderates the relationship between team faultlines and performance. In particular, we specify the condition under which group faultline affects team success negatively: when the faultline-induced subgroups do not communicate with each other on an above-average frequency.

Taken together, to understand collective outcomes in complex systems, including success in teams, we need to investigate those processes that constitute the system's complexity instead of merely studying the characteristics of its elements. Although we found that cross-subgroup communication does not predict success itself, we demonstrated that it is a critical variable in those complex interdependencies that influence performance. Our finding on the specifying role of communication is consistent with prior research that reported on communication as a moderating factor in the gender diversity-performance relationship [120]. Moreover, to the best of our knowledge, the current study is the first that operationalizes communication as real-time interactions during problem solving providing a truly process variable, and it examines it in the faultline-performance relationship. Furthermore, we accounted for the structure and time-varying nature of these communication ties recognizing that communication is not a static and homogeneous phenomenon, but it is often distributed across time and team members in an asymmetric manner.

Knowing when communication is more or less critical for teams can be vital information for organizations that regularly rely on teamwork to enable them to maximize performance. Considering team composition, a factor that can be easily manipulated by managers, a possible implication of faultline theory is to prevent the formation of these dividing lines. This could be achieved by composing rather homogeneous groups based on one or two diversity dimensions, for instance, on gender or ethnicity – an undesirable goal that would deepen social inequalities especially regarding the access to job opportunities for minority groups. Furthermore, this may distort the supply-demand relations of the job market by regulating who should be selected for certain positions where applicants' qualifications and competence would be less decisive. Therefore, instead of preventing faultline formation, a more realistic and desirable way is to understand the mechanisms by which faultlines can influence team outcomes and manage them. From this perspective, our result showing that team-wide, real-time communication mitigates faultline effect is an important insight. This finding suggests that managers should facilitate communication between the faultline-induced subgroups as much as possible to ensure efficient team functioning for tasks that require the synthetized knowledge of the whole team as a

single unit. At the same time, strong faultline-groups with low cross-subgroup communication might be assigned to tasks that can be easily divided and later integrated into their final forms by leaders or coordinators. Both strong- and weak-faultline groups are present in real life, and can function efficiently, if we understand their core mechanisms and know how to build on the strength of these teams.

Our novel research design utilizing escape rooms could also provide far-reaching potential for practical applications. Although, in this work, we were interested in teams as adaptive task performing systems, this environment could be used for evaluating individual social behavior as well, similarly to assessment centers. As opposed to psychological tests that can be easily manipulated, escape rooms considered as non-interventional social laboratories can portray a prospective team member's wide range of characteristics including their socioemotional dynamics, participation level, behavioral answers to time pressure, attitudes toward other members, etc. Therefore, escape rooms could be employed already during the selection process, and not only for team buildings. Although there is a great variety of tasks in assessment centers, escape rooms can present a treasure chest of information on social behavior during a typically longer duration, in a more controlled but less stressful environment of a selection process compared to assessment centers.

Nevertheless, regarding the present work, we are cautious on a generalizability of our results. A clear limitation of our data is its small, but not unprecedentedly small sample size. This is related to the time-effort-and resource consuming collection of real-time interaction data despite the advantages of escape rooms over laboratory experiments in this regard as well. In addition, the employed survival models consider the temporal nature of communication, and as a byproduct of this method, we observed time-group level data including much more data points than what we had on the team level ignoring communication dynamics. Although, the performed statistical analysis accounts for sample size, and thus, it does not discount the validity of our findings based on significant relationships, we should be careful with rejecting unsupported hypothesis. Last, but not least, although teams in escape rooms meet the definition of project teams in terms of the non-routine nature of the task, we are aware that in many cases, project teams are composed of members whose skills complement each other. However, in contrast to the previous future directions, this matter could only be addressed by significant intervention in the research setting, as one would have to recruit members and design the experiment rigorously, which could lessen the advantages of the present data collection such as its non-interventional nature and the minimized Hawthorne effect.
In addition to the indisputably desirable bigger sample, future research could

also provide a more sophisticated picture on the information components of communication, whether communication is directed to the task or social relationships, for example. Moreover, several control variables could be incorporated into future research such as controlling for potentially varying levels of task motivation and extraversion of team members – two individual-level characteristics that might influence the amount of cross-subgroup communication. Items measuring these individual aspects could be included in the questionnaire on socio-demographic traits.

In short, our work contributes to faultline research by investigating team familiarity and communication between faultline-induced subgroups as predictors and moderators of team success. In line with the theoretical underpinnings of CEM and the network perspective, examining team processes is crucial to capture the diversity-performance relationship in its richness. Based on this idea and by relying on the innovative quasi-experimental setting of escape rooms, we incorporated real process variables into our analysis. We found that real-time communication (during non-routine problem-solving) across the faultline-induced subgroups plays a crucial role to understand when diversity can affect success in collaborative problem-solving. We hope that our work encourages more research analyzing teams in their complexity.

CHAPTER 4

THE ANATOMY OF SOCIAL DYNAMICS IN ESCAPE ROOMS: TEMPORAL, SIGNED AND GROUP INTERACTIONS

4.1 Prelude

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 to others [2]. Social interactions are indeed key to understand information exchange [78] and social contagion [145, 146, 147, 148]. 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 contexts [45, 149, 47, 46, 41, 44, 37], drawing attention to the importance of high-frequency time-resolved social processes [48] to understand collective intelligence [39] and collaborative problem solving [40, 41, 42, 43]. In the last decades, network science has proved to be a powerful and flexible framework to understand the complex relational structure of human dynamics [150, 151], from structural balance theory [152, 153] to the detection of emergent mesoscale structures such as communities [154] and cores [155], associated with the coordinated behavior of multiple individuals in human societies.

This chapter portrays a nuanced and meticulous picture about the social dynamics of problem-solving teams in escape rooms. Exploiting the innovative setting of escape rooms and the collected fine-grained data, we analyze teams' high-resolution social dynamics, including collaborative network evolution and

conversation rules guiding communication, and 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. This chapter highlights the potential of escape rooms to investigate task-performing teams in a minimally biased environment, contributing to advance the new science of teams.

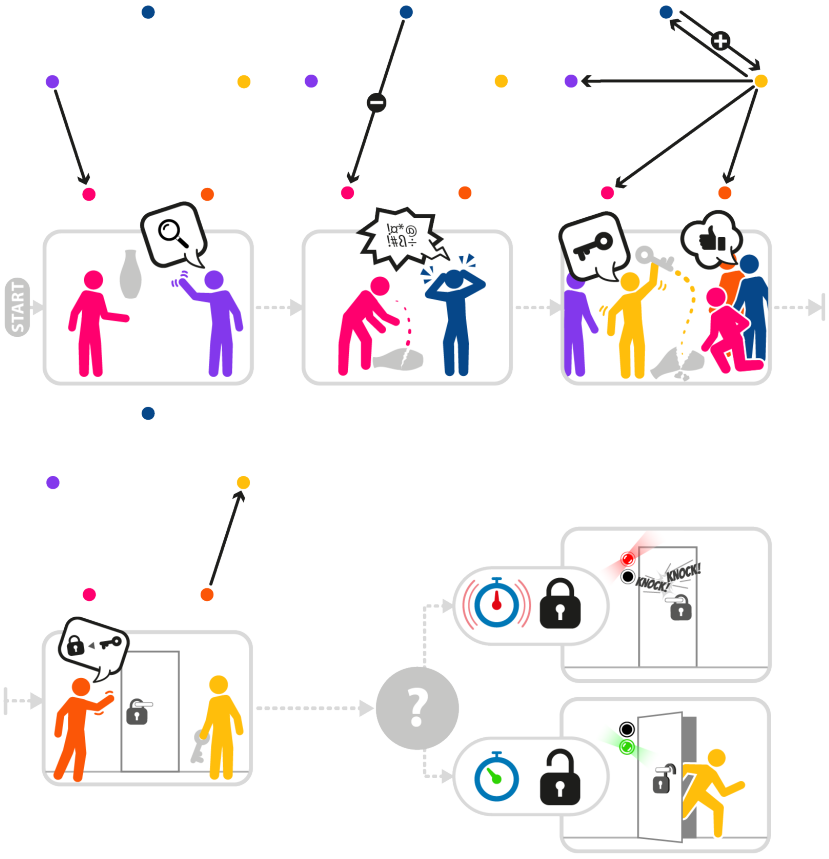


Figure 4.1. ***The complex and diverse dimensions of social interactions in escape rooms.*** *Interactions are represented as temporal networks, which capture skeletal structure of communication between team members. Each interaction is directed, can be emotionally neutral or charged (either positively or negatively), addressed to a particular person (pairwise) or to a wider group.*

As detailed in Chapter 2, for each team (composed of 4 or 5 individuals, for a total of 171 players, all from Hungary) we extracted high-resolution temporal interactions from video records. Again, we deemed teams that managed to escape within one hour 'successful', while teams unable to do so were labeled as 'failed' groups. As a brief reminder, teams in escape rooms, must explore and exploit information, possibly through collaborations, by searching for clues, deciphering codes or opening locks. 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. Moreover, using this questionnaire, we asked participants to appoint the leader of their group if there is any.

Our data can be mapped as a temporal network, where interactions among individuals occur at specific points in time [156]. In Fig. 4.1 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). In the following, we provide a data-driven characterization of different dimensions of the social dynamics taking place in escape rooms.

4.2 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 seconds.

We find that older and more educated players speak for longer stretches at a time (Figs. 4.2A,B). The strength of a relationship between individuals has a significant role in determining the intensity of their collaborative pairwise interactions during problem-solving. In particular, prior member familiarity promotes communication, and the more time two players have known each other, the higher their rate of communications (Fig. 4.3A). 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. 4.3B)

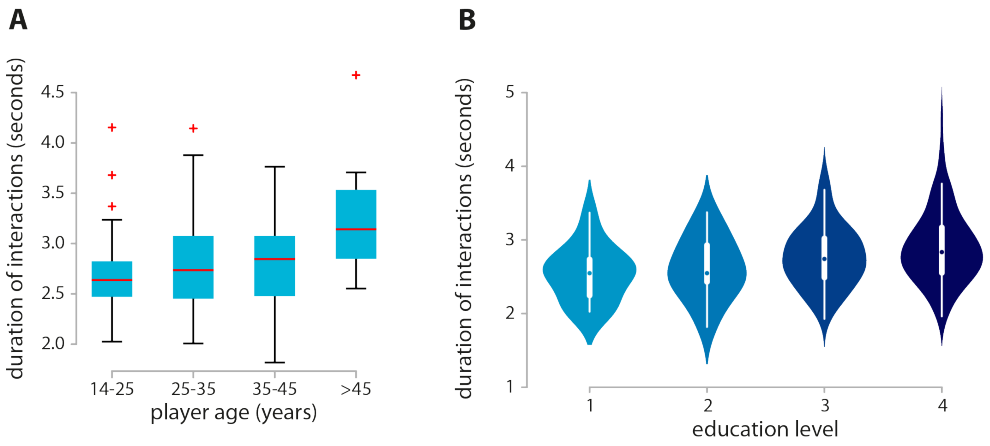


Figure 4.2. ***Interactions and demographics.*** *Individual demographics of team members determine their interaction patterns. Older (A) and more educated (B) members tend to initiate longer interactions.*

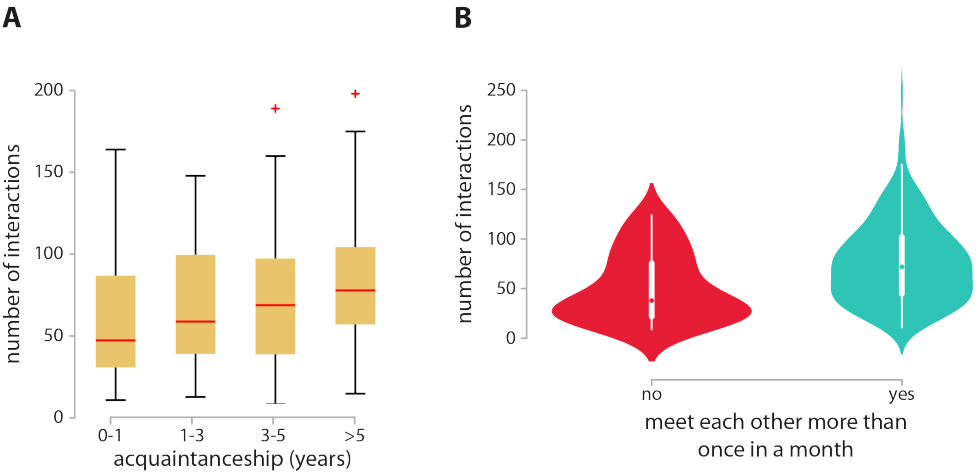


Figure 4.3. ***Interactions and group relations.*** *Relational characteristics of members determine their interaction patterns. Both the length of prior acquaintanceship (A) and meeting frequency (B) are associated with an increased number of interactions during problem-solving.*

4.3 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 Gibson [1]. 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.

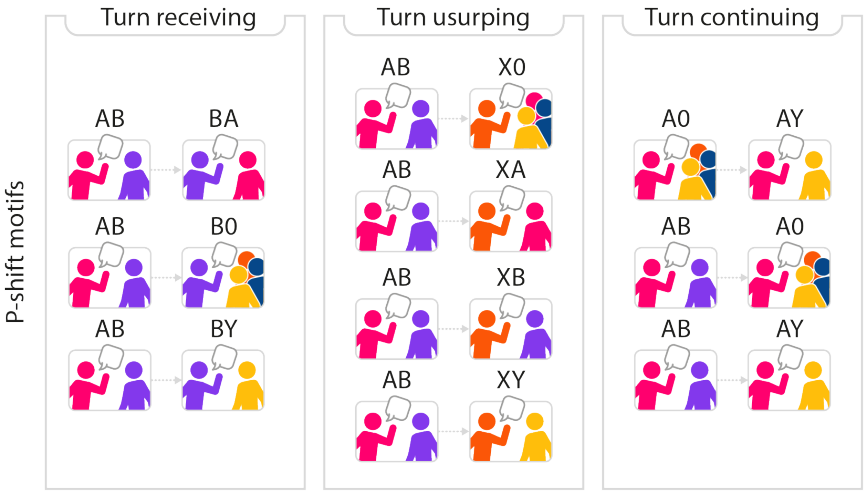


Figure 4.4. ***Inventory of Participation Shift Profiles.*** *We identify ordered communication sequences in escape rooms by Gibson's Participation Shift Profile framework [1]. P-shift motifs started with a directed remark (pairwise) can be clubbed up into 3 categories denoting the way by which the second speaker get they turn.*

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. 4.4), on which we elaborate in the following. By comparing them with what found in a suitable null model, we determine whether the motif frequencies of the empirical data are significantly

different from those observed by random chance in systems which preserve the same number of interactions, but where temporal correlations are washed away (Fig. 4.5).

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.

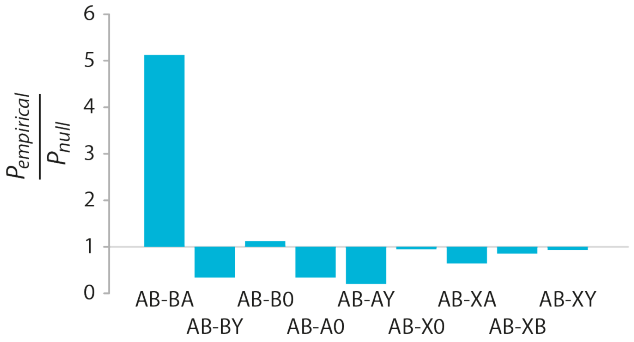


Figure 4.5. ***Conversation rules in escape rooms.*** *The prototypical turn-receiving P-shift AB-BA (which refers to an immediate reciprocation by B is overrepresented at the expense of AB-BY, while pairwise turn-continuing motifs such as AB-AY and AB-A0 are underrepresented in escape rooms.*

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 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. As Gibson predicted based on his case study [1], "holders of formal authority had a penchant for turn-continuing". Here, we show that his statement originally developed in the context of man-

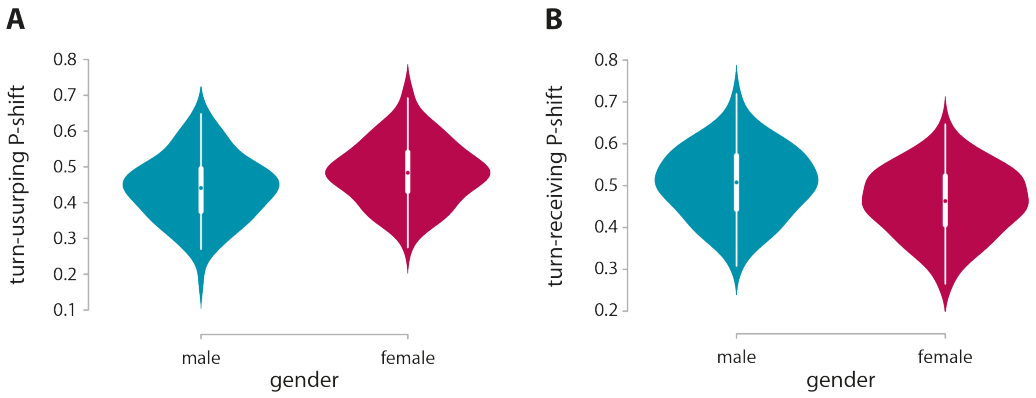


Figure 4.6. ***P-shifts and gender.*** *Men and women participate differently in team conversations. Women are more likely to usurp the turn (A) as opposed to men. This is probably connected to the observation that women tend to receive less opportunity to speak (B). Thus, they have to seize the floor to express their opinion.*

agerial teams with diverse goals, it also holds true for problem-solving project teams. Although the presence of these P-shifts is the rarest, we found that, the pre-established leader - a team member who has been nominated by at least 50% of the team as a leader - inclines to emerge in turn-continuing motifs (Fig. 4.5/C). Moreover, the higher level of authorization (number of nominations) they receive from team members, the stronger they tendency is to turn continue ($r_s = 0.41$, $p_{val} = 0.043$) and the weaker the inclination is to speak after they received the turn as B in motif AB-BA ($r_s = -0.38$, $p_{val} = 0.058$).

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* [1]. 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. 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. 4.6) 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. 4.6). 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 observations [157].

4.4 Signed interactions

Our experiment also gave us access to information about the emotional load of each interactions, classified as neutral, positive or negative. Praising, encouragement, and more 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 [51, 52, 53].

We find that both older (Fig. 4.7) 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-demograhic 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. 4.8 suggests that these emotional interactions are not equally distributed. Successful teams are more balanced in terms of emotional

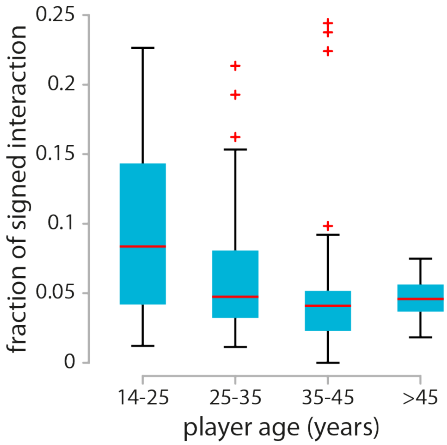


Figure 4.7. ***Signed interactions and age.*** *Older players initiate less signed communication than younger members, suggesting a greater focus on task performance.*

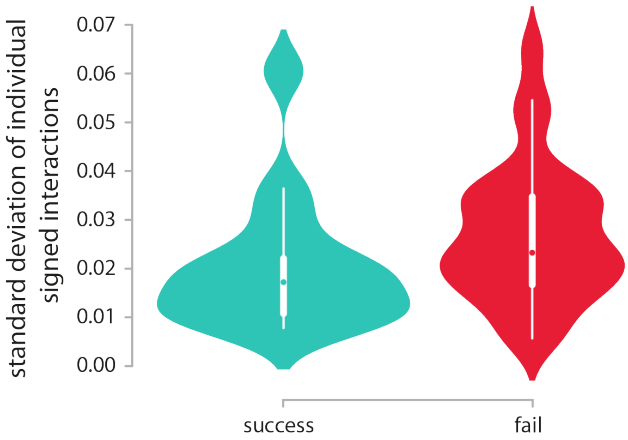


Figure 4.8. ***Emotional balance and team success.*** *While the total amount of emotionally charged interactions is similar in successful and failed teams, successful groups are more emotionally balanced across their individuals than the failed ones. This is quantified by the standard deviations of signed interactions across team members.*

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. 4.9A, 5-minute rolling window applied). 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.

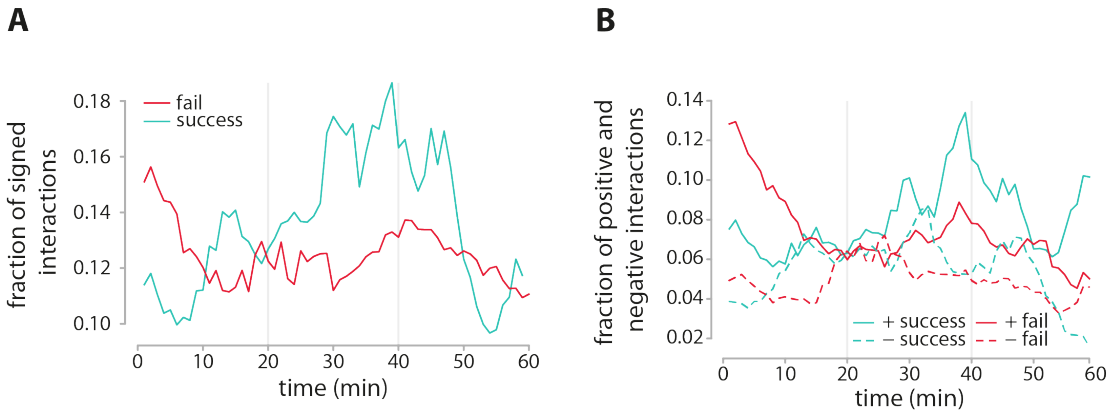


Figure 4.9. ***Emotional dynamics and team success.*** *Failed teams are characterized by a greater frequency of positive interactions in the initial stage of the game, which however rapidly declines over time (A). This trend is largely due to the high initial frequency of positive ties (B), possibly 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.*

We can gain more insights on these patterns by decomposing signed interactions into positive and negative ties (Fig. 4.9B, 5-minute rolling window applied). Surprisingly, the initially greater amount of emotionally loaded interactions in failed teams is due to a high rate of positive interactions (13%, doubling the amount of successful teams), which however rapidly declines within the first 15 minutes. This unexpected feature might reflect a lower focus (e.g. making jokes, laughing), 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 minutes, 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 minute mark, only the remaining successful teams contribute to the curves shown in Fig. 4.9A,B.

4.5 Group interactions

Beyond pairwise communication, we also investigate group interactions [158], where more than a single recipient is addressed. Group interactions constitute $\approx$ 20% of all the interactions, and are typically 0.77 shorter than one-to-one interactions. We find that older (Fig. 4.10A) and more educated (Fig. 4.10B) 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.

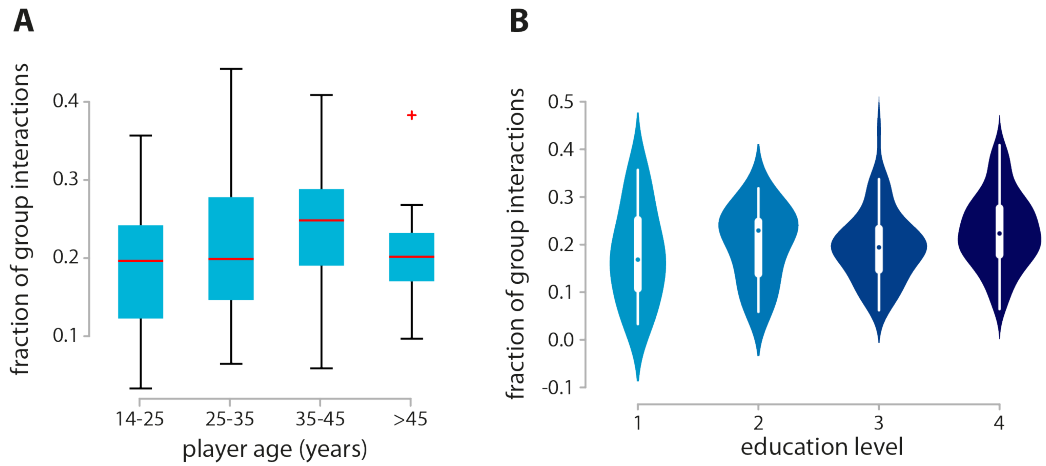


Figure 4.10. ***Group interactions and demographics.*** *Older (A) and more educated (B) players display a higher frequency of group interactions, suggesting an unequal level of confidence and authority in the teams.*

Observing the temporal dynamics (in 5-minute rolling windows, Fig. 4.11),

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 minutes, 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 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. 4.11.

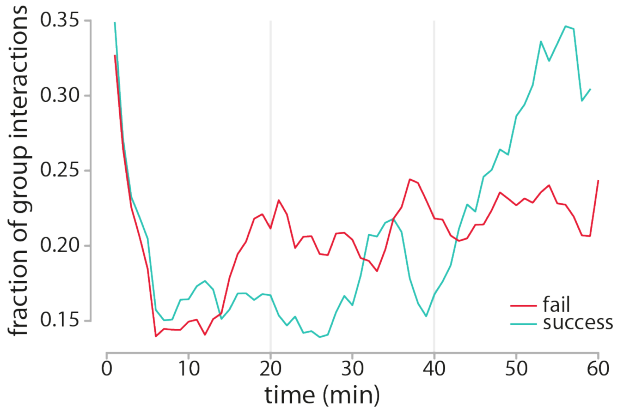


Figure 4.11. ***Temporal group interactions and team success.*** *An investigation of the dynamics of group interactions reveal that after an initial get-together, members tend to work in smaller groups. An increase in group interactions for failed teams around the 20-minute mark reflects the emergence of possible early issues, while the rapid increase for successful teams in the latest part of the game is associated with celebration.*

Conversation rules governing group interactions can be analysed by a suitable extension of the previously introduced P-shift profiles [1]. Following a group remark, where *0* identifies the group, either the speaker continues their turn (as *A* in *A0-AY*, turn continuing), or someone else claims the turn (turn

claiming). Turn-claiming can happen in three distinct ways, as *X* in *A0-XA/A0-X0/A0-XY*, where the addressee can be the first speaker *A*, the group *0*, 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 *A0-X0* or *A0-XA* at the expense of *A0-XY* (Fig. 4.12). These two P-shifts overrepresented in our data cover those situations when a remark addressed to the group is followed by a reaction.

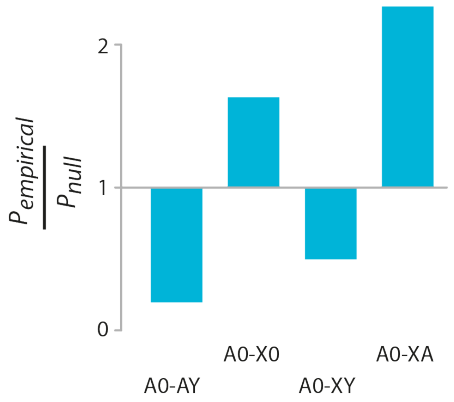


Figure 4.12. ***Group interactions and P-shifts.*** *P-shift analysis reveals the overabundance of motifs A0-X0 and A0-XA, typically associated with the delivery of complex messages.*

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. 4.12, no significant difference between successful and failed teams). We find that the most overabundant motif is *A0-XA*, the 'group version' of the previously discussed *AB-BA* pairwise motif [1]. A second group motif, *A0-X0*, 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 (*A0-X0*), or asks for clarification to *A*'s action on behalf of the team (*A0-XA*).

4.6 Leadership and authority

Leadership is a cooperative process that cannot be understood independently of followership [159, 160, 161]. Therefore, leadership – as one of the most salient social roles – is an outspoken indicator of group hierarchy. Leadership refers to "guidance of others in their pursuits, often by organizing, directing, coordinating, supporting, and motivating their efforts [84](pp.246)".

Our leadership data come from two different sources. On the one hand, we acquired information on who is the leader of the teams (if any) by asking the members to name a person in the questionnaire (pre-established leader). On the other hand, we measured leadership relying on the interaction data. Here, we intended to detect classic, centralized leadership implying control and giving orders. Actors who possess high out-degree centrality are considered as opinion leaders, which refers to the extent to which a team member is able to influence other members' attitudes or overt behaviour in a desired way informally and frequently[49]. Thus, we identified the highest out-degree holders, that is, the players with the highest participation level during problem-solving as opinion leaders. We compared this emergent leader with the pre-established leader. In those cases, when there is a a pre-established leader of the team, and this person is identical to the emergent leader, it is assumed that teams have a clear perception on their intragroup roles. In the opposite case, when the pre-established leader is not the opinion leader during problem-solving, teams either have a less obvious vertical structure of roles, or the initial hierarchy can be dissolved during the game. The third option is when there is no leader of the team due to either the lack of consensus or because the majority of votes went to nobody.

We find that teams with a mismatch between the pre-established and emergent leader are efficient in 55%, while none of the groups with a clear perception on intragroup roles managed to exit the room in the given timeframe. Escape room teams with no leader proved to be successful in 30% (Fig. 4.13). In other words, teams with a clear perception on the leader of the groups (match), that is, those groups having a rather rigid role structure with some level of hierarchy, failed to complete the task uniformly. In contrast, teams with an unidentifiable leader, and thus, supposedly with a rather flexible role structure or a dissoluble hierarchy (mismatch) were efficient in a much higher percent - with even higher success rate than those teams with admittedly no leader. This result might reflect on the the fact that the environment team performing the task under time pressure is relatively random and unpredictable for each member regardless of the different amount of authority they might possess.

In teams with a rather rigid role structure, there is a higher chance for a

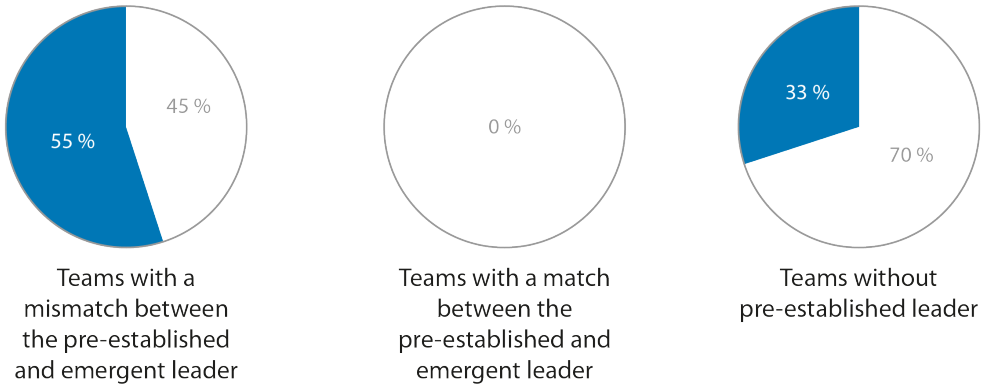


Figure 4.13. ***Leadership and success rate.*** *Comparing the pre-established leader with the emergent leader, we investigated if the fact whether teams have or do not have a clear perception of their leader (match between the pre-established and the emergent leader) facilitates team success. We found that teams without a clear perception of their leader tend to be more successful, suggesting that a similar amount of intra-group authority enhances performance in problem-solving.*

group culture dominated by implicit expectations connected to behavioral constraints and respect for authority. Members of these groups might be reluctant to express their half-finished thoughts but think well what and when they speak. While, teams with a flexible role structure or less diversity in the amount of members' authority, we assume a more elevated atmosphere where idea sharing is likely to be smoother facilitating efficient group work.

4.7 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 may 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. Pre-established leaders, as holders of formal authority were found to have a tendency to continue their speaking turns. 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 prosperous 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. Interestingly, successful teams can be also characterized by a more fluid role structure than that of failed groups, suggesting that groups with less hierarchy and more agility tend to do better in escape rooms.

In summary, here we have proposed escape rooms as an innovative research setting for studying groups in 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 this 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.

CHAPTER 5

CONCLUSION

This thesis demonstrated the utility of escape rooms in researching human interactions. In Chapter 2, I introduced escape rooms as minimally biased, non-interventional social laboratories free from the typical weaknesses of both laboratory experiments and field studies. I collected real-time, fine-grained data on the collaboration network of small teams during problem solving in escape rooms relying on video records. In addition, I obtained teams' demographic and relational information by a questionnaire designed to explore and relate sociodemographic characteristics to social dynamics. This unique and novel quasi-experimental setting allowed me to study the collaborative problem solving activity of intact, non-manipulated teams in a meticulous manner. With this research design, I also intended to contribute to the increasingly rare studies of interacting groups that carry out actual activity as opposed to just thinking about something, to study human interactions based on a direct and objective data source.

Specifically, leveraging this unique data enabled us to extend faultline research by studying - for the first time - the role of collaborative communication in the faultline-success relationship, where communication is measured by real-time interaction data. The main finding of Chapter 3 suggests that when investigating team success, the effect of diversity and collaborative communication can only be understood in relation to each other. We found that group diversity is detrimental to team success *when* the faultline-induced subgroups do not interact with each other on an above-average frequency. This suggests that managers should facilitate communication between the faultline-induced subgroups as much as possible to ensure an efficient team functioning for task that requires the synthesized knowledge of the whole team as a single unit. On the other hand, strong faultline-groups with low cross-subgroup communica-

tion might be assigned to tasks that can be easily divided and later integrated into their final forms by leaders. In a more abstract sense, our finding reinforces a fundamental concept of the network approach to investigate those processes that constitute a system's complexity instead of merely studying the characteristics of its elements.

A more data-driven, case study in Chapter 4 provided a nuanced picture about teams' social dynamics and interpersonal behavior in escape rooms. We treated interactions as temporal networks, which captured the skeletal structure of communication between team members. We have identified peculiar interaction patterns (e.g. longer and typically more task-directed communication often addressed to the group) linked to specific demographic characteristics (older and more educated team members). We have also seen that the length of interpersonal relationships contributes to active problem-solving interactions, and that the dominant conversation rules [1] are manifested in action-reaction pair motifs reflecting the high demand for practical actions in escape rooms. Confirming previous studies on women's disadvantage [157], we also found that women need to assert themselves to seize the floor to speak in order to express their opinion as opposed to men, who typically receive the turn to speak.
Considering team success, our analysis revealed different patterns for successful and failed groups in two dimensions of interactions. First, investigating signed networks, successful teams were found to be more engaged in task-directed communication at the beginning of the task performance. At later stages, though relationship-oriented communication seemed to have a positive influence on efficiency. Second, successful groups typically had a more fragmented temporal collaboration network than failed teams. However, we identified group-wide conversation around the middle and the end of the game when members gathered to synthesize their knowledge - a potential sign of a conscious strategy.
Finally, we found that successful teams tend to have a rather similar amount of authority which presumably enables them to create an elevated group atmosphere entailing smooth and efficient communication in this setting.

Understanding teams in their complexity requires nuanced and detailed information about their social dynamics that accounts for the core process of any team-related emergent phenomenon. This thesis introduced escape rooms as an innovative research setting for studying groups in controlled, yet non-manipulated environment. The utilization of this unique, high resolution data can pave the way for both research aiming to address complex theoretical questions or to explore the data itself systematically to gain new insights into human problem solving behavior. I argue that using escape rooms as easily accessible social laboratories carries the potential to open up new directions

at the intersection of sociology, organization science, network science and small group research by contributing to a deeper understanding of intra-group processes and collective action. This thesis is the first step into this direction.

My data still has numerous untapped opportunities. One of these potential research directions we are currently addressing is marked by a task-based inquiry. Throughout this thesis, the unit of observation were teams, and their success was operationalized as a binary variable (escaped/non-escaped). At the same time, in escape rooms, several sub-tasks have to be implemented by the teams to get the opportunity to face the ultimate task, that is, to reach the final door that leads out of the room. Proposing a different perspective for our future research avenue, we go to the level of these sub-tasks. Following this problem-based approach, on the one hand, we can work with an increased amount of data (number of teams times the number of sub-tasks implemented by them). On the other hand, we can use the problem solving time as a more sophisticated measure of team success. Consequently, we can utilize the high temporal resolution of the data even more comprehensively.
The sub-tasks in both escape rooms have a peculiar arrangement that shows the dependency between them (Figure 5.1). All sub-tasks can be categorized as either exploration or exploitation. Exploration refers to the act of searching for clues and relevant information pieces, while exploitation is about utilizing and combining knowledge. Thus, implementing an exploitation task assumes that the relevant information for that has already been explored. This arrangement of escape room tasks structures problem-solving time naturally. We use these time slices to investigate team success in the following way. First, we calculate how much time it took for the team to implement the sub-task after all related conditions were met - our new measure for success. Then, we seek to identify behavioral patterns predicting problem-solving time in these time intervals assuming that efficient problem solving at exploration phases might require different collaborative activity than that of at exploitation phases.

Although this work is still in its infancy, we can already see two promising indicators, the average number of connected component of collaboration networks, and the standard deviation of out-degrees. Our preliminary findings show that teams with a high average number of (second-based) connected components are faster in solving exploration tasks, but not when implementing exploitation tasks. Moreover, teams performing both exploration and exploitation tasks are also faster when members speak on a similar frequency. These initial results suggest that multitasking (simultaneous operation of sub-teams on different sub-tasks) enhance efficient performance of exploration task, but not that of exploitation tasks. In general, the higher differences are between team

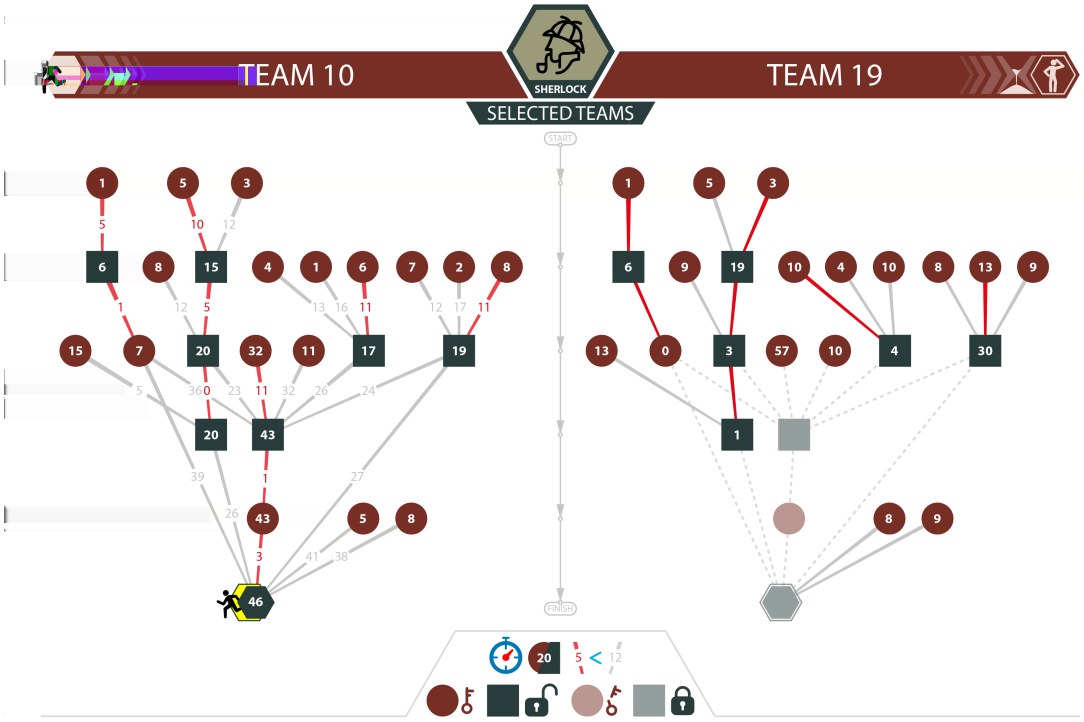


Figure 5.1. ***Task dependency in room Sherlock*** *presented on the example of a successful (Team 10) and a failed (Team 19) team. We can see that Team 10 accomplished all tasks and was able to exit the room in minute 46 as opposed to Team 19 that failed to solve many tasks. Circles represent exploration phases, while squares denote exploitation tasks with the numbers showing the absolute time elapsed to perform the given task. In general, the edges indicate the time difference between two nodes, and the red ties are those with minimum time spent between two dependent tasks.*

members in the amount of initiated communication ties, the more time it takes for them to implement either an exploration or an exploitation task. Therefore, egalitarianism – hereby referred to as a similar amount of interaction initiated by each team member – seems to assist the progress of collaborative problem solving. Enriching this theme, we also plan to investigate if there is an optimal organization of sub-task implementation.

Regarding long-term plans, stochastic actor-based models (SAOM), and exponential random graph models (ERGM) present themselves as natural methods to investigate the temporal escape room data among many other potential ways to further explore traits of collaborative problem solving. We hope that this thesis sparks interest in these directions.

CHAPTER 6

APPENDICES

6.0.1 Table A1

Predictors	**Model A5** *Log-Odds*	**Model A6** *Log-Odds*	**Model A7** *Log-Odds*	**Model A8** *Log-Odds*
(Intercept)	-9.94^{**}	-9.98^{**}	-9.17^{**}	-11.64^{**}
	(3.66)	(3.77)	(3.57)	(3.86)
time (min)	0.23^{***}	0.23^{***}	0.23^{***}	0.25^{***}
	(0.05)	(0.05)	(0.06)	(0.06)
average age	0.02	0.00	0.03	0.02
	(0.08)	(0.09)	(0.08)	(0.09)
age diversity	-0.27	-0.28	-0.36	-0.34
	(0.30)	(0.24)	(0.21)	(0.24)
education diversity	-4.30^{**}	-2.12	-4.86^{**}	-3.63^{*}
	(1.67)	(2.56)	(1.79)	(1.63)
gender diversity	-4.53^{*}	-4.14^{*}	-6.20^{*}	-5.58^{**}
	(1.86)	(1.80)	(2.94)	(1.97)
faultline strength	-3.05	-2.70	-3.79	-0.64
	(1.97)	(2.00)	(2.05)	(2.27)
room	0.21	0.09	0.02	0.26
	(0.75)	(0.64)	(0.63)	(0.65)
average acquaintanceship	-0.25	-0.20	-0.33^{*}	-0.20
	(0.13)	(0.12)	(0.14)	(0.12)
average acquaintanceship *age diversity	-0.01 (0.05)			
average acquaintanceship *education diversity		-0.51 (0.49)		
average acquaintanceship *gender diversity			0.29 (0.38)	
average acquaintanceship *faultline strength				-0.38 (0.24)
Observations	2308	2308	2308	2308
R^2 Tjur	0.096	0.099	0.098	0.114
		$^{*}p < 0.05$	$^{**}p < 0.01$	$^{***}p < 0.001$

Table 6.1. *Discrete-time survival models on exiting the room - models with interactions.*

6.0.2 Consent form

CONSENT FORM

Study Title: The micro-dynamic nature of group interactions

What is this study about?

The purpose of the study is to reveal how project teams' collaboration network evolve across time and team members, and how it correlates with successfulness.

What will I do if I choose to be in this study?

You will be asked to give your permission to us to use the video data on your group's activity (game) that is recorded by the escape room operators. This is the data from which we can construct the team's collaboration network and analyse its dynamics in terms of task performance.
Also, you will be asked to fill a 5-minute questionnaire to inform us about some demographics and group relations.
No sensitive confidential data will be requested. A random ID/chosen name by you will be used to protect your identity, (unless you specifically request that you be identified by your true name).

How will you protect the information you collect about me, and how will that information be shared?

Results of this study may be used in publications and presentations. We will not use any sensitive confidential data, as we will not collect identifiable personal information.

Who can I contact if I have questions?

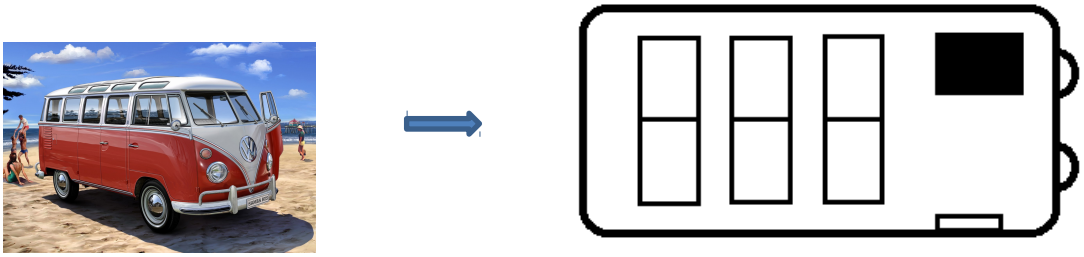
If you have questions, you are free to ask them now. If you have questions later, you may contact the researcher(s) at Szabo_Rebeka@phd.ceu.edu

6.0.3 Questionnaire

Dear Participant,

This questionnaire is part of my data collection for my PhD research, in which I observe the temporal changes of group interactions and its correlation to task-solving.
At the item(s) of the questionnaire requiring names, the indication of the first name or nickname is sufficient and only serves to distinguish team members.
Estimated time of the questionnaire: **5 minutes**.

1. Imagine that you go for a one-day bus trip with your current team-mates. The figure below represents the seats on the bus. Please have your team mates "sat down" including yourself. *(One square represents one seat, therefore one name is supposed to get there.)*



2. Is there a leader of the group among your current team-mates?
(If there is, please, write down his/her name and that why you consider him/her as the leader.)

.....

3. How long have you known each other with your current team mates?
(Please answer this question in case of each of your team-mates.)

a) Name (first name/nickname):

.....

I know him/her (approx.) .....year/month ago;
From (school/workplace/via a friend/relative/relationship, etc.):.....;
We talk, meet: *each day / each week / biweekly / monthly / less often*

b) Name (first name/nickname):

.....

I know him/her (approx.) .....year/month ago;
From (school/workplace/via a friend/relative/relationship, etc.).....;
We talk, meet: *each day / each week / biweekly / monthly / less often*

c) Name (first name/nickname):

.....

I know him/her (approx.) .....year/month ago;
From (school/workplace/via a friend/relative/relationship, etc.):.....;
We talk, meet: *each day / each week / biweekly / monthly / less often*

d) Name (first name/nickname):

.....

I know him/her (approx.) .....year/month ago;
From (school/workplace/via a friend/relative/relationship, etc.):.....;
We talk, meet: *each day / each week / biweekly / monthly / less often*

e) Name (first name/nickname):

.....

I know him/her (approx.) .....year/month ago;
From (school/workplace/via a friend/relative/relationship, etc.):.....;
We talk, meet: *each day / each week / biweekly / monthly / less often*

4. Your current job: .....

5. Highest level of education: .....

6. Your age: .....

7. Your first name/nickname (what you have used before in the questionnaire):

CONSENT

I have read the Consent Form and the research study has been explained to me. I agree to participate in this research study.

.....
Participant's first name/nicname

....
Participant's signiture Date

Thank you for your help!

Have fun!

6.0.4 Transcription guide

TRANSCRIPTION GUIDE

Quality aspects:

1. If you are **not sure about something**, anywhere (e.g.: who talks to whom, what the emotional charge of the communication is, etc.). Please, **ask!** - we'll discuss and work it out together! This is very important for quality!

2. If you feel **your brain is running out of your ears** and every part of your body, both physically and mentally, is protesting against continuing the coding, let me know! **Relax** a little and rather have the description later than be inaccurate and put fatigue at the expense of precision!

WhatsApp/telefon: O. Szabó Rebeka, +36703832223, or after 7 pm: o.sz.rebu@gmail.com

Technical aspects:

Coding scheme - example

start time	end time	sender	receiver	charge	transcription
21:32:01	21:32:05	E	M	pos	text
21:32:01	21:32:05	E	D	pos	text
21:32:01	21:32:05	E	A	pos	text
21:32:06	22:32:09	D	A	neg	text
22:32:09	23:32:10	A	M	neutral	text
23:32:10	23:32:15	D	M	neg	text
---------- Cash register ----------					
21:32:12	21:32:16	M	D	pos	text

!!!IMPORTANT!!!

- An **interaction** is more than communication! For example, if a player shows something to another player without speaking, it is also considered an interaction, so we record it!

- Other non-verbal communication can be nodding; 'reply' for a request for example giving something to someone upon asking, etc. In contrast, natural moving in the room while searching for clues is not an interaction.

- An interaction is one train of thought. So, if someone starts talking but takes a tiny break (takes a breath – max. 1-2 seconds) and then continues while there is no other interaction in the meantime, it can be taken as one piece of interaction.

To be recorded:

1. start time
2. end time
3. sender
4. receiver
5. emotional charge
6. transcription
7. general impression about the team
8. sub-results

1. Start time:

Record the second in which the interaction starts.

2. End time:

The last second in which the interaction is ongoing. I f somebody speaks until 21:32:09, so that in the 9th second they do not speak, the proper end time to be recorded is 21:32:08 as this is the last second the interaction is happening.

3-4. Sender-receiver:

- The sender and receiver/addressee of the interaction.
- If someone speaks to several people at the same time (a group interaction), we record in the same way as in the example above: E speaks to ME, D, and MA at the same time, because all three lines have the same start and end time of the interaction.

+ operator's call: we record when the interaction starts and ends, and who speaks with the operator

5. Emotional charge:

We classify the emotional charge of interactions into 3 categories: positive, negative and neutral.

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.***

Examples:

- praise,
- laughing,
- joking,
- inspiration/emotional support, etc.

That is every interaction that influence team spirit positively.

- Smooth collaboration between team members on a task does NOT account for positive interactions until these interactions are not emotionally loaded.
- Example: A: ‘we have to smell these tobaccos and identify them’ → B: ‘wow, cool!’: neutral, as B’s inspiration does not reflect to A’s action (but to the situation/circumstances)

Negative interactions *are all interactions that relate to the previous action negatively/****have negative emotional charge AND is not directed to the task itself but is rather related to interpersonal relationships.***

Examples:

- disapproving,
- despising,
- trivialize,
- ironic, cynical reaction,
- intentional ignoration of someone
- quarrel, etc.

That is every interaction that influence team spirit negatively.

- The word ‘no’ does not make the interaction automatically negative!
- When one draws the other’s attention to a fact that the other has thought wrongly and happens to have the word ‘no’ in it, the interaction is still neutral because the point is to pass on facts / communicate information.
- E.g.: A: ‘here is a mongoose’ → B: ‘no, because look, it is the pipe…’. So, the interaction is tension-free, AND it is not a conflict of individual opinions, but a statement of fact (=task-related interaction=neutral). It would be negative if someone raises, they voice and there is a touchable tension.

Neutral interactions** are **directed to the task itself and not to the relationships AND are not emotionally loaded.

Examples:

- finding of facts,
- emotionless communication,
- questions,
- showing something, etc.

SPECIAL CASES:

1. **Cutting into one's sentence while talking to another person,** e.g. A cuts into ZS's sentence in a raised voice while (s)he talks to T**.** This is an important interaction that should be coded as two interactions: 1) A talks to T (neutral); 2) A send a negative interaction to ZS (as A silences ZS).

2. Approval, but nervous reaction: negative – so the **strong emotional charge overwrites the emotional charge** of the interaction. From neutral to negative in this case.

3. **Talking to herself/himself OR any action which is not shared by someone:** So it is not addressed to anyone, not a reaction to a previous action, also when it is so quiet that nobody could have a chance to hear it => both the sender and the receiver are the same person, e.g.: A talk to A (as (s)he grumbles something under her/his nose)

4. If there are parts with **different emotional charge within the same sentence**, it is coded in two separate lines with the corresponding emotional charge (this is very rare!)
   E.g.: 'I think it's a pig, so you said pure nonsense first'. - although the first half of the sentence is a statement and has a neutral emotional charge, it is used by the speaker to justify the second part of the sentence (that the addressee is stupid). So, it is organically connected, so overall it is one negative interaction.
   (It could be separated if there is another interaction during the sentence since we are coding in chronological order.)

Another example:

A: ‘we can't even unload this’

- if (s)he grumbles under her/his nose, (s)he is the sender and the receiver
- if (s)he says it out loud, it is a derogatory (neg.) sentence addressed to everyone, and it is easy to see that it can be destructive to team spirit

6. Transciption:

Accurate description of communication! - Any questions, you don't hear something, please, put a question mark, I'll look into it later.

The text description should be as complete as possible! This includes the audible part of the conversation with the operator! (+ non-verbal interactions, of course)

7. General impression on team spirit:

- as a team: lively / enthusiastic / innervated, etc.?
- as players: Everyone is equally motivated? Some are very enthusiastic, while others are quieter and more withdrawn?

8. Sub-results:

If the subtask is not fulfilled entirely, for example in the case of the first code, if the coin is placed in the wallet but the 4 cards are not found, the partial result should not be recorded as it is not considered a valid subresult.

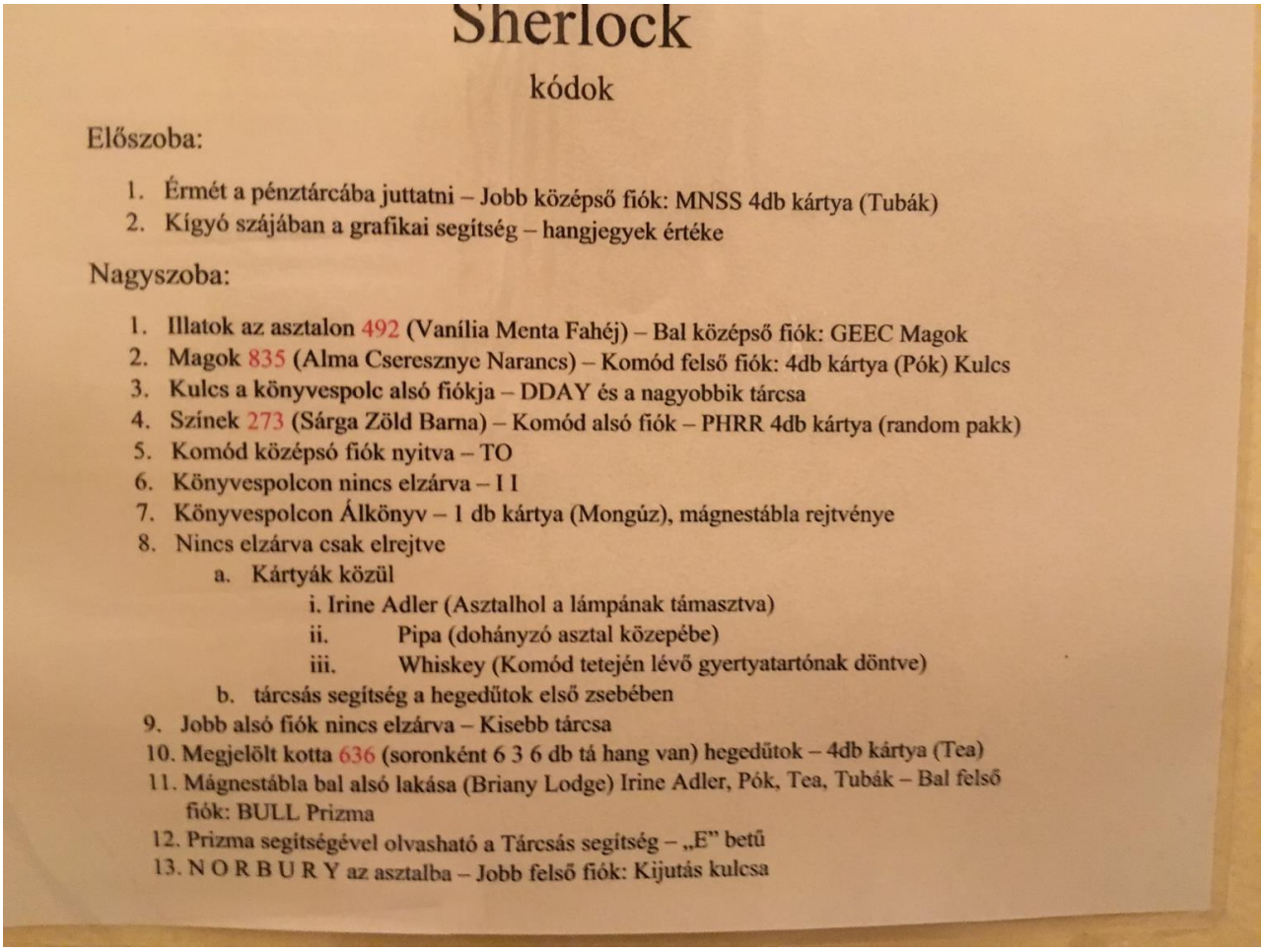
Only a fully solved subtask can be considered as completed, partial result.

For the exact subresult codes, please see the flowchart of the rooms and use the provided names of the subresults uniformly.

AT THE END:

The game & coding is over when the key of the door is found and identified for what it is for.
(Sometimes they don't know what it opens for minutes.)

HELP FOR CODING SUBTASKS



Sherlock

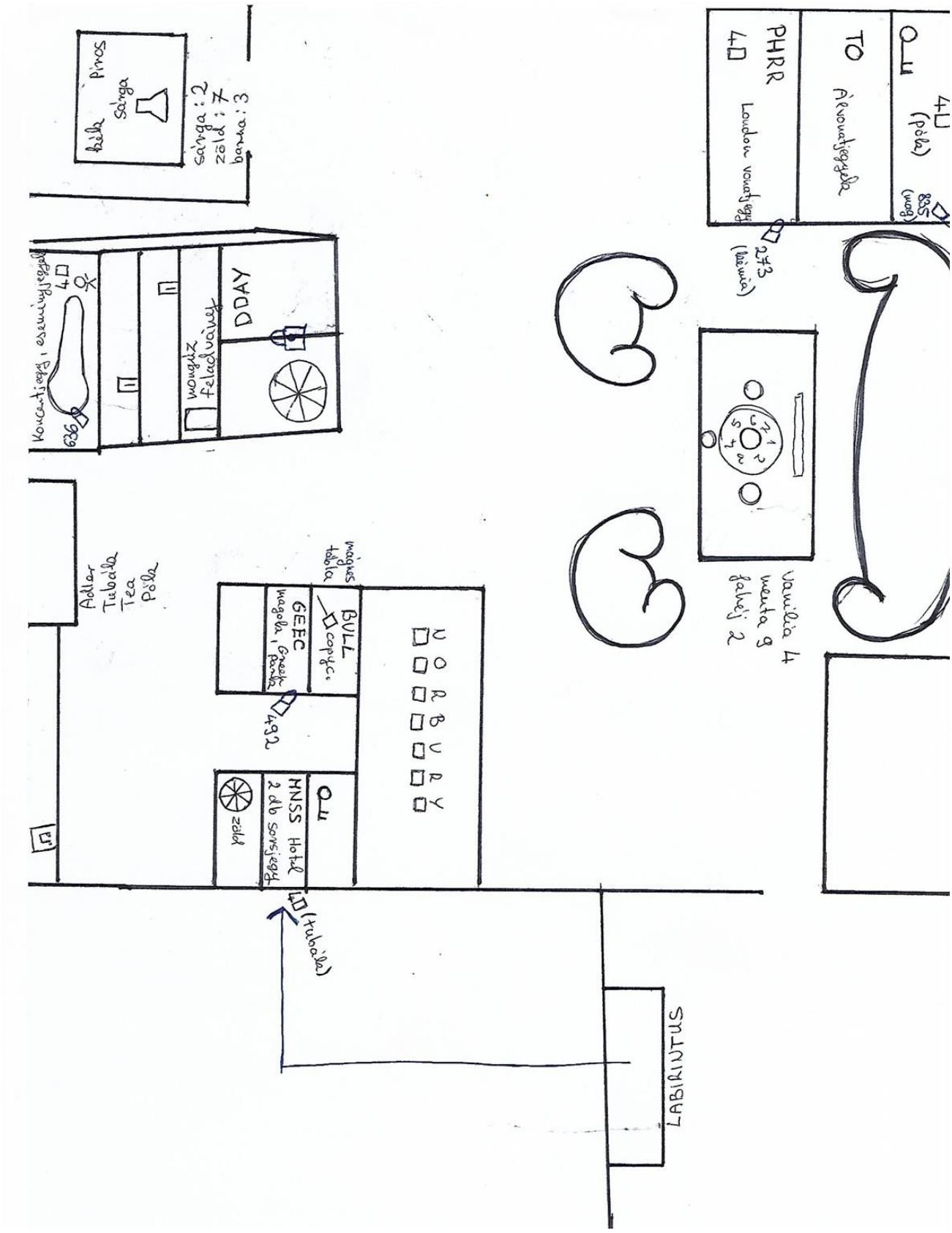
kódok

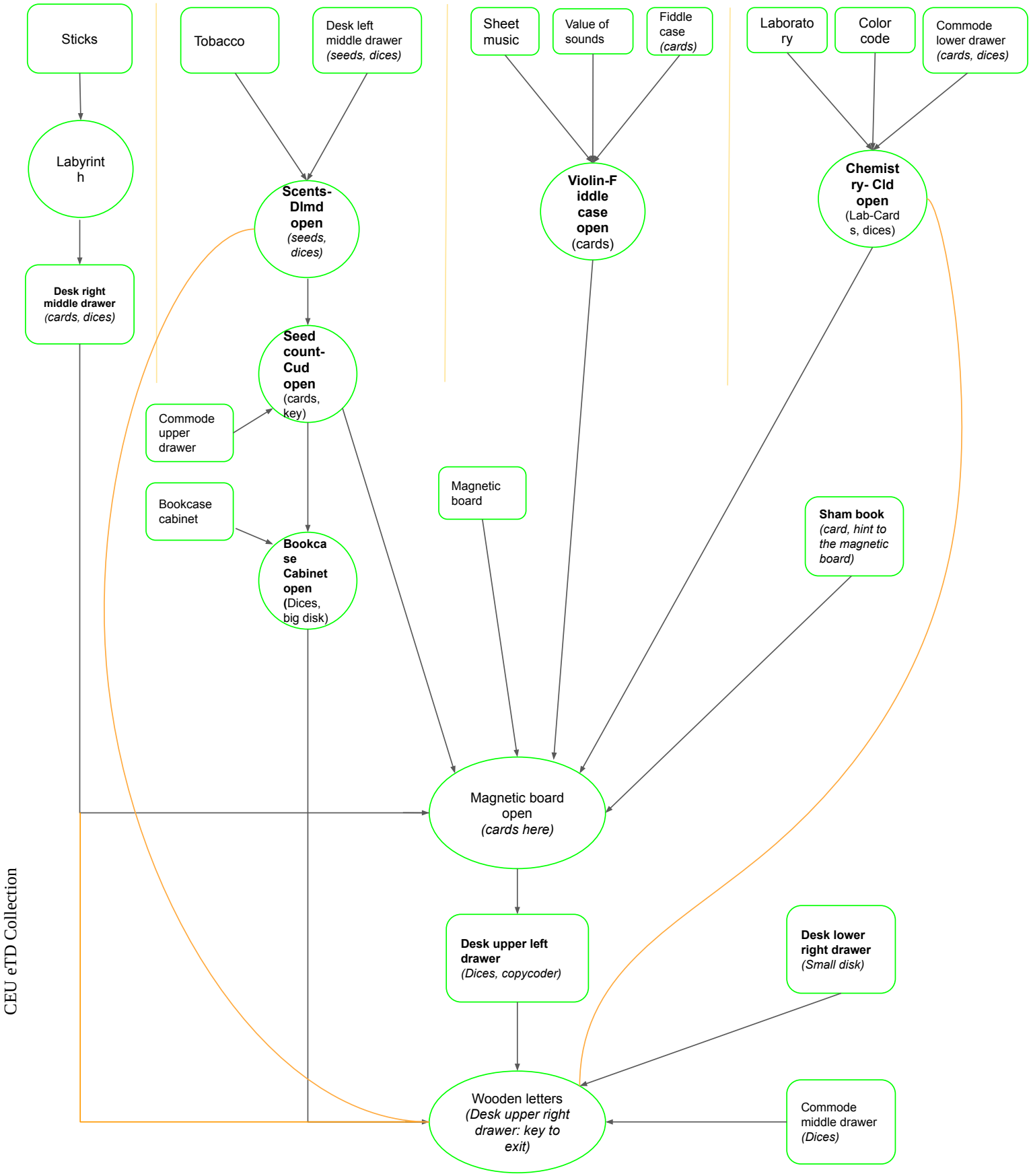
Előszoba:

1. Érmét a pénztárcába juttatni – Jobb középső fiók: MNSS 4db kártya (Tubák)
2. Kígyó szájában a grafikai segítség – hangjegyek értéke

Nagyszoba:

1. Illatok az asztalon 492 (Vanília Menta Fahéj) – Bal középső fiók: GEEC Magok
2. Magok 835 (Alma Cseresznye Narancs) – Komód felső fiók: 4db kártya (Pók) Kulcs
3. Kulcs a könyvespolc alsó fiókja – DDAY és a nagyobbik tárcsa
4. Színek 273 (Sárga Zöld Barna) – Komód alsó fiók – PHRR 4db kártya (random pakk)
5. Komód középső fiók nyitva – TO
6. Könyvespolcon nincs elzárva – I I
7. Könyvespolcon Álkönyv – 1 db kártya (Mongúz), mágnestábla rejtvénye
8. Nincs elzárva csak elrejtve
   a. Kártyák közül
      i. Irine Adler (Asztalhol a lámpának támasztva)
      ii. Pipa (dohányzó asztal közepébe)
      iii. Whiskey (Komód tetején lévő gyertyatartónak döntve)
   b. tárcsás segítség a hegedűtok első zsebében
9. Jobb alsó fiók nincs elzárva – Kisebb tárcsa
10. Megjelölt kotta 636 (soronként 6 3 6 db tá hang van) hegedűtok – 4db kártya (Tea)
11. Mágnestábla bal alsó lakása (Briany Lodge) Irine Adler, Pók, Tea, Tubák – Bal felső fiók: BULL Prizma
12. Prizma segítségével olvasható a Tárcsás segítség – „E" betű
13. N O R B U R Y az asztalba – Jobb felső fiók: Kijutás kulcsa



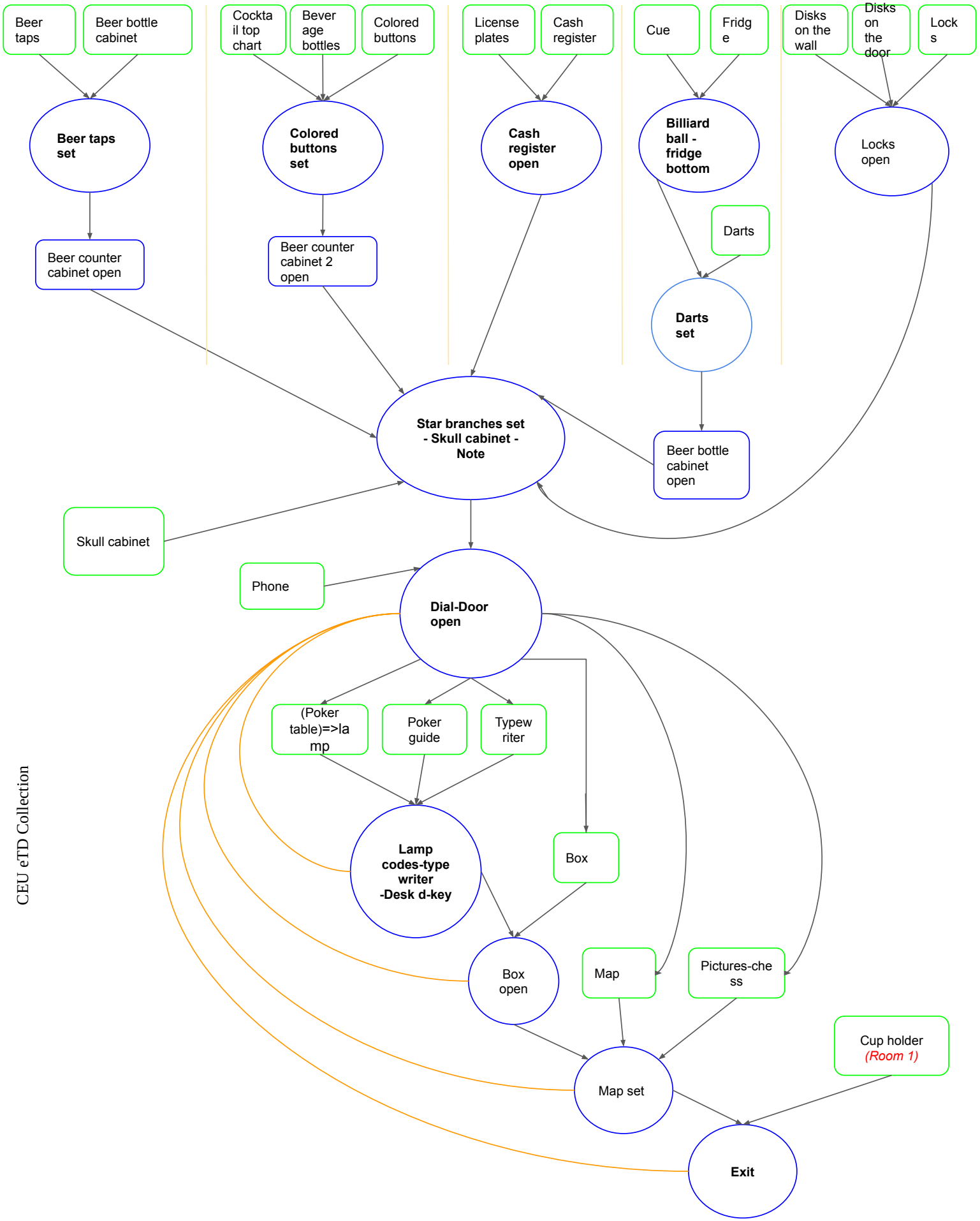


Exploration: first occasion an actor becomes aware of an exploration space

Exploitation: a (knowledge) combination task has been implemented properly

Arrows: task dependency; All tasks arrowed to the next depending phase

KERESZTAPA



BIBLIOGRAPHY

[1] D. R. Gibson, "Participation shifts: Order and differentiation in group conversation," *Social forces*, vol. 81, no. 4, pp. 1335–1380, 2003.

[2] B. Uzzi, "Embeddedness in the making of financial capital: How social relations and networks benefit firms seeking financing," *American sociological review*, pp. 481–505, 1999.

[3] H. A. Simon, "The structure of ill structured problems," *Artificial intelligence*, vol. 4, no. 3-4, pp. 181–201, 1973.

[4] J. Bruggeman, "Consensus, cohesion and connectivity," *Social Networks*, vol. 52, pp. 115–119, 2018.

[5] G. R. VandenBos, *APA dictionary of psychology.* American Psychological Association, 2007.

[6] J. H. Fewell, D. Armbruster, J. Ingraham, A. Petersen, and J. S. Waters, "Basketball teams as strategic networks," *PloS one*, vol. 7, no. 11, p. e47445, 2012.

[7] R. Guimera, B. Uzzi, J. Spiro, and L. A. N. Amaral, "Team assembly mechanisms determine collaboration network structure and team performance," *Science*, vol. 308, no. 5722, pp. 697–702, 2005.

[8] S. Wuchty, B. F. Jones, and B. Uzzi, "The increasing dominance of teams in production of knowledge," *Science*, vol. 316, no. 5827, pp. 1036–1039, 2007.

[9] L. Hunter and E. Leahey, "Collaborative research in sociology: Trends and contributing factors," *The American Sociologist*, vol. 39, no. 4, pp. 290–306, 2008.

[10] B. F. Jones, S. Wuchty, and B. Uzzi, "Multi-university research teams: Shifting impact, geography, and stratification in science," *science*, vol. 322, no. 5905, pp. 1259–1262, 2008.

[11] S. Torrisi, S. Manfredi, I. Iacopini, and V. Latora, "Creative connectivity project-a network based approach to understand correlations between interdisciplinary group dynamics and creative performance," *E&PDE 2019 Towards a New Innovation Lanscape*, pp. 530–535, 2019.

[12] S. E. Jackson, A. Joshi, and N. L. Erhardt, "Recent research on team and organizational diversity: Swot analysis and implications," *Journal of management*, vol. 29, no. 6, pp. 801–830, 2003.

[13] A. Zeng, Y. Fan, Z. Di, Y. Wang, and S. Havlin, "Fresh teams are associated with original and multidisciplinary research," *Nature Human Behaviour*, pp. 1–9, 2021.

[14] B. Uzzi, S. Wuchty, J. Spiro, and B. F. Jones, "Scientific teams and networks change the face of knowledge creation," in *Networks in Social Policy Problems*, pp. 47–59, Cambridge University Press, 2012.

[15] D. A. Harrison, S. Mohammed, J. E. McGrath, A. T. Florey, and S. W. Vanderstoep, "Time matters in team performance: Effects of member familiarity, entrainment, and task discontinuity on speed and quality," *Personnel Psychology*, vol. 56, no. 3, pp. 633–669, 2003.

[16] A. M. Petersen, "Quantifying the impact of weak, strong, and super ties in scientific careers," *Proceedings of the National Academy of Sciences*, vol. 112, no. 34, pp. E4671–E4680, 2015.

[17] S. Mukherjee, Y. Huang, J. Neidhardt, B. Uzzi, and N. Contractor, "Prior shared success predicts victory in team competitions," *Nature human behaviour*, vol. 3, no. 1, pp. 74–81, 2019.

[18] L. Wu, D. Wang, and J. A. Evans, "Large teams develop and small teams disrupt science and technology," *Nature*, vol. 566, no. 7744, pp. 378–382, 2019.

[19] K. Börner, N. Contractor, H. J. Falk-Krzesinski, S. M. Fiore, K. L. Hall, J. Keyton, B. Spring, D. Stokols, W. Trochim, and B. Uzzi, "A multi-level systems perspective for the science of team science," *Science Translational Medicine*, vol. 2, no. 49, pp. 49cm24–49cm24, 2010.

[20] B. Uzzi, S. Mukherjee, M. Stringer, and B. Jones, "Atypical combinations and scientific impact," *Science*, vol. 342, no. 6157, pp. 468–472, 2013.

[21] J. S. Coleman, "Social capital in the creation of human capital," *American journal of sociology*, vol. 94, pp. S95–S120, 1988.

[22] G. C. Homans, "The human group. new brunswick," 1950.

[23] A. Bavelas, "Communication patterns in task-oriented groups," *The journal of the acoustical society of America*, vol. 22, no. 6, pp. 725–730, 1950.

[24] A. Bavelas, D. Barrett, A. M. Association, *et al.*, "An experimental approach to organizational communication. publications (massachusetts institute of technology. dept. of economics and social science)," *Industrial Relations. American Management Association*, 1951.

[25] H. J. Leavitt, "Some effects of certain communication patterns on group performance.," *The Journal of Abnormal and Social Psychology*, vol. 46, no. 1, p. 38, 1951.

[26] R. T. Sparrowe, R. C. Liden, S. J. Wayne, and M. L. Kraimer, "Social networks and the performance of individuals and groups," *Academy of management journal*, vol. 44, no. 2, pp. 316–325, 2001.

[27] T. T. Baldwin, M. D. Bedell, and J. L. Johnson, "The social fabric of a team-based mba program: Network effects on student satisfaction and performance," *Academy of management journal*, vol. 40, no. 6, pp. 1369–1397, 1997.

[28] R. Reagans and E. W. Zuckerman, "Networks, diversity, and productivity: The social capital of corporate r&d teams," *Organization science*, vol. 12, no. 4, pp. 502–517, 2001.

[29] B. Uzzi, "The sources and consequences of embeddedness for the economic performance of organizations: The network effect," *American sociological review*, pp. 674–698, 1996.

[30] H. A. Landsberger, "Hawthorne revisited: Management and the worker, its critics, and developments in human relations in industry.," 1958.

[31] R. T. A. Leenders, N. S. Contractor, and L. A. DeChurch, "Once upon a time: Understanding team processes as relational event networks," *Organizational Psychology Review*, vol. 6, no. 1, pp. 92–115, 2016.

[32] S. W. Kozlowski and K. J. Klein, "A multilevel approach to theory and research in organizations: Contextual, temporal, and emergent processes.," 2000.

[33] R. Y. Hirokawa and A. J. Salazar, "Task-group communication and decision-making performance," *The handbook of group communication theory and research*, pp. 167–191, 1999.

[34] N. Katz, D. Lazer, H. Arrow, and N. Contractor, "Network theory and small groups," *Small group research*, vol. 35, no. 3, pp. 307–332, 2004.

[35] J. E. McGrath and L. Argote, "Group processes in organizational contexts," *Blackwell handbook of social psychology: Group processes*, pp. 603–627, 2001.

[36] C. Downs, "Clampitt, pg, & pfeffer, a.(1988). communication and organizational out—comes," *Handbook of organizational communica-tion*, pp. 171–211.

[37] P. Sapiezynski, A. Stopczynski, D. D. Lassen, and S. Lehmann, "Interaction data from the copenhagen networks study," *Scientific Data*, vol. 6, no. 1, pp. 1–10, 2019.

[38] J. N. Cummings and R. Cross, "Structural properties of work groups and their consequences for performance," *Social networks*, vol. 25, no. 3, pp. 197–210, 2003.

[39] A. W. Woolley, C. F. Chabris, A. Pentland, N. Hashmi, and T. W. Malone, "Evidence for a collective intelligence factor in the performance of human groups," *science*, vol. 330, no. 6004, pp. 686–688, 2010.

[40] Y.-A. De Montjoye, A. Stopczynski, E. Shmueli, A. Pentland, and S. Lehmann, "The strength of the strongest ties in collaborative problem solving," *Scientific reports*, vol. 4, no. 1, pp. 1–6, 2014.

[41] V. Sekara, A. Stopczynski, and S. Lehmann, "Fundamental structures of dynamic social networks," *Proceedings of the national academy of sciences*, vol. 113, no. 36, pp. 9977–9982, 2016.

[42] C. J. Gomez and D. M. Lazer, "Clustering knowledge and dispersing abilities enhances collective problem solving in a network," *Nature communications*, vol. 10, no. 1, pp. 1–11, 2019.

[43] B. Monechi, G. Pullano, and V. Loreto, "Efficient team structures in an open-ended cooperative creativity experiment," *Proceedings of the National Academy of Sciences*, vol. 116, no. 44, pp. 22088–22093, 2019.

[44] O. Lederman, D. Calacci, A. MacMullen, D. C. Fehder, F. E. Murray, and A. Pentland, "Open badges: A low-cost toolkit for measuring team communication and dynamics," *arXiv preprint arXiv:1710.01842*, 2017.

[45] A. Pentland, "Socially aware, computation and communication," *Computer*, vol. 38, no. 3, pp. 33–40, 2005.

[46] G. Callahan, "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)," *The Review of Austrian Economics*, vol. 29, no. 1, pp. 93–97, 2016.

[47] A. Pentland and T. Heibeck, "Honest signals," *MIT press*, 2008.

[48] A. Pentland, "The new science of building great teams," *Harvard business review*, vol. 90, no. 4, pp. 60–69, 2012.

[49] E. Rogers, "Diffusion of innovations, 5th edn tampa," *FL: Free Press.[Google Scholar]*, 2003.

[50] J. G. Adair, "The hawthorne effect: a reconsideration of the methodological artifact.," *Journal of applied psychology*, vol. 69, no. 2, p. 334, 1984.

[51] K. A. Jehn, "A qualitative analysis of conflict types and dimensions in organizational groups," *Administrative science quarterly*, pp. 530–557, 1997.

[52] T. L. Simons and R. S. Peterson, "Task conflict and relationship conflict in top management teams: the pivotal role of intragroup trust.," *Journal of applied psychology*, vol. 85, no. 1, p. 102, 2000.

[53] C. K. De Dreu and L. R. Weingart, "Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis.," *Journal of applied Psychology*, vol. 88, no. 4, p. 741, 2003.

[54] K. Krippendorff, "Computing krippendorff's alpha-reliability," 2011.

[55] K. Krippendorff, "Reliability in content analysis: Some common misconceptions and recommendations," *Human communication research*, vol. 30, no. 3, pp. 411–433, 2004.

[56] L. Sproull, S. Kiesler, and S. B. Kiesler, *Connections: New ways of working in the networked organization*. MIT press, 1991.

[57] J. E. McGrath, "Time, interaction, and performance (tip) a theory of groups," *Small group research*, vol. 22, no. 2, pp. 147–174, 1991.

[58] S. G. Cohen and D. E. Bailey, "What makes teams work: Group effectiveness research from the shop floor to the executive suite," *Journal of management*, vol. 23, no. 3, pp. 239–290, 1997.

[59] D. A. Harrison and K. J. Klein, "What's the difference? diversity constructs as separation, variety, or disparity in organizations," *Academy of management review*, vol. 32, no. 4, pp. 1199–1228, 2007.

[60] R. Reagans, E. Zuckerman, and B. McEvily, "How to make the team: Social networks vs. demography as criteria for designing effective teams," *Administrative science quarterly*, vol. 49, no. 1, pp. 101–133, 2004.

[61] T. H. Cox, S. A. Lobel, and P. L. McLeod, "Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task," *Academy of management journal*, vol. 34, no. 4, pp. 827–847, 1991.

[62] C. A. Easley, "Developing valuing and managing diversity in the new millennium," *Organization Development Journal*, vol. 19, no. 4, p. 38, 2001.

[63] J. C. Turner, "Social categorization and the self-concept: A social cognitive theory of group behavior.," 2010.

[64] S. M. Thatcher and P. C. Patel, "Group faultlines: A review, integration, and guide to future research," *Journal of Management*, vol. 38, no. 4, pp. 969–1009, 2012.

[65] D. C. Lau and J. K. Murnighan, "Demographic diversity and faultlines: The compositional dynamics of organizational groups," *Academy of management review*, vol. 23, no. 2, pp. 325–340, 1998.

[66] D. Van Knippenberg and M. C. Schippers, "Work group diversity," *Annu. Rev. Psychol.*, vol. 58, pp. 515–541, 2007.

[67] M. A. Cronin, K. Bezrukova, L. R. Weingart, and C. H. Tinsley, "Subgroups within a team: The role of cognitive and affective integration," *Journal of Organizational Behavior*, vol. 32, no. 6, pp. 831–849, 2011.

[68] M. J. Pearsall, A. P. Ellis, and J. M. Evans, "Unlocking the effects of gender faultlines on team creativity: Is activation the key?," *Journal of Applied Psychology*, vol. 93, no. 1, p. 225, 2008.

[69] J. T. Polzer, C. B. Crisp, S. L. Jarvenpaa, and J. W. Kim, "Extending the faultline model to geographically dispersed teams: How colocated subgroups can impair group functioning," *Academy of management Journal*, vol. 49, no. 4, pp. 679–692, 2006.

[70] C. Gibson and F. Vermeulen, "A healthy divide: Subgroups as a stimulus for team learning behavior," *Administrative science quarterly*, vol. 48, no. 2, pp. 202–239, 2003.

[71] D. C. Lau and J. K. Murnighan, "Interactions within groups and subgroups: The effects of demographic faultlines," *Academy of management journal*, vol. 48, no. 4, pp. 645–659, 2005.

[72] K. Bezrukova, K. A. Jehn, E. L. Zanutto, and S. M. Thatcher, "Do workgroup faultlines help or hurt? a moderated model of faultlines, team identification, and group performance," *Organization science*, vol. 20, no. 1, pp. 35–50, 2009.

[73] K. Bezrukova, C. S. Spell, and J. L. Perry, "Violent splits or healthy divides? coping with injustice through faultlines," *Personnel Psychology*, vol. 63, no. 3, pp. 719–751, 2010.

[74] R. Reagans, L. Argote, and D. Brooks, "Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together," *Management science*, vol. 51, no. 6, pp. 869–881, 2005.

[75] J. A. Espinosa, S. A. Slaughter, R. E. Kraut, and J. D. Herbsleb, "Familiarity, complexity, and team performance in geographically distributed software development," *Organization science*, vol. 18, no. 4, pp. 613–630, 2007.

[76] R. S. Huckman, B. R. Staats, and D. M. Upton, "Team familiarity, role experience, and performance: Evidence from indian software services," *Management science*, vol. 55, no. 1, pp. 85–100, 2009.

[77] R. S. Huckman and B. R. Staats, "Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance," *Manufacturing & Service Operations Management*, vol. 13, no. 3, pp. 310–328, 2011.

[78] M. S. Granovetter, "The strength of weak ties," *American journal of sociology*, vol. 78, no. 6, pp. 1360–1380, 1973.

[79] P. V. Marsden and K. E. Campbell, "Reflections on conceptualizing and measuring tie strength," *Social forces*, vol. 91, no. 1, pp. 17–23, 2012.

[80] D. Melamed and B. Simpson, "Strong ties promote the evolution of cooperation in dynamic networks," *Social networks*, vol. 45, pp. 32–44, 2016.

[81] J. M. Levine and E. R. Smith, "Group cognition: Collective information search and distribution.," 2013.

[82] D. Vora and L. Markóczy, "Group learning and performance: The role of communication and faultlines," *The International Journal of Human Resource Management*, vol. 23, no. 11, pp. 2374–2392, 2012.

[83] J. Lipman-Blumen, H. J. Leavitt, *et al.*, *Hot groups: Seeding them, feeding them, and using them to ignite your organization*. Oxford University Press on Demand, 1999.

[84] D. R. Forsyth, *Group dynamics*. Cengage Learning, 2018.

[85] C. H. Loch and Y. Wu, *Behavioral operations management*. Now Publishers Inc, 2007.

[86] F. Gino and G. Pisano, "Toward a theory of behavioral operations," *Manufacturing & Service Operations Management*, vol. 10, no. 4, pp. 676–691, 2008.

[87] E. R. Crawford and J. A. LePine, "A configural theory of team processes: Accounting for the structure of taskwork and teamwork," *Academy of Management Review*, vol. 38, no. 1, pp. 32–48, 2013.

[88] D. J. Devine, L. D. Clayton, J. L. Philips, B. B. Dunford, and S. B. Melner, "Teams in organizations: Prevalence, characteristics, and effectiveness," *Small group research*, vol. 30, no. 6, pp. 678–711, 1999.

[89] J. E. Mathieu, J. R. Hollenbeck, D. van Knippenberg, and D. R. Ilgen, "A century of work teams in the journal of applied psychology.," *Journal of applied psychology*, vol. 102, no. 3, p. 452, 2017.

[90] T. Rickards and S. Moger, "Creative leadership processes in project team development: an alternative to tuckman's stage model," *British journal of Management*, vol. 11, no. 4, pp. 273–283, 2000.

[91] M. E. M. Barak and D. J. Travis, "Socioeconomic trends: Broadening the diversity ecosystem," in *The Oxford handbook of diversity and work*, p. 393, Oxford University Press, 2012.

[92] D. Van Knippenberg and J. N. Mell, "Past, present, and potential future of team diversity research: From compositional diversity to emergent diversity," *Organizational Behavior and Human Decision Processes*, vol. 136, pp. 135–145, 2016.

[93] F. J. Milliken and L. L. Martins, "Searching for common threads: Understanding the multiple effects of diversity in organizational groups," *Academy of management review*, vol. 21, no. 2, pp. 402–433, 1996.

[94] H. Van Dijk, M. van Engen, and D. van Knippenberg, "Work group diversity and performance: A meta-analysis," in *Academy of Management Annual Meeting, Chicago, IL*, 2009.

[95] D. Van Knippenberg, J. F. Dawson, M. A. West, and A. C. Homan, "Diversity faultlines, shared objectives, and top management team performance," *human relations*, vol. 64, no. 3, pp. 307–336, 2011.

[96] D. R. Ilgen, J. R. Hollenbeck, M. Johnson, and D. Jundt, "Teams in organizations: From input-process-output models to imoi models," *Annu. Rev. Psychol.*, vol. 56, pp. 517–543, 2005.

[97] D. Van Knippenberg, C. K. De Dreu, and A. C. Homan, "Work group diversity and group performance: an integrative model and research agenda.," *Journal of applied psychology*, vol. 89, no. 6, p. 1008, 2004.

[98] H. Tajfel and J. Turner, "The social identity theory of intergroup behavior. in s. worchel and wg austing (eds.). psychology of intergroup relations," 1986.

[99] K. Y. Williams and C. A. O'Reilly III, "Demography and," *Research in organizational behavior*, vol. 20, pp. 77–140, 1998.

[100] J. Li and D. C. Hambrick, "Factional groups: A new vantage on demographic faultlines, conflict, and disintegration in work teams," *Academy of Management Journal*, vol. 48, no. 5, pp. 794–813, 2005.

[101] E. Mannix and M. A. Neale, "What differences make a difference? the promise and reality of diverse teams in organizations," *Psychological science in the public interest*, vol. 6, no. 2, pp. 31–55, 2005.

[102] I. L. Janis, "Victims of groupthink: A psychological study of foreign-policy decisions and fiascoes.," 1972.

[103] K. A. Bantel and S. E. Jackson, "Top management and innovations in banking: Does the composition of the top team make a difference?," *Strategic management journal*, vol. 10, no. S1, pp. 107–124, 1989.

[104] C. R. Østergaard, B. Timmermans, and K. Kristinsson, "Does a different view create something new? the effect of employee diversity on innovation," *Research policy*, vol. 40, no. 3, pp. 500–509, 2011.

[105] S. T. Bell, A. J. Villado, M. A. Lukasik, L. Belau, and A. L. Briggs, "Getting specific about demographic diversity variable and team performance relationships: A meta-analysis," *Journal of management*, vol. 37, no. 3, pp. 709–743, 2011.

[106] B. Meyer, M. Shemla, J. Li, and J. Wegge, "On the same side of the fault-line: Inclusion in the leader's subgroup and employee performance," *Journal of Management Studies*, vol. 52, no. 3, pp. 354–380, 2015.

[107] A. C. Homan, D. Van Knippenberg, G. A. Van Kleef, and C. K. De Dreu, "Bridging faultlines by valuing diversity: diversity beliefs, information elaboration, and performance in diverse work groups.," *Journal of applied psychology*, vol. 92, no. 5, p. 1189, 2007.

[108] A. C. Homan, J. R. Hollenbeck, S. E. Humphrey, D. V. Knippenberg, D. R. Ilgen, and G. A. Van Kleef, "Facing differences with an open mind: Openness to experience, salience of intragroup differences, and performance of diverse work groups," *Academy of Management Journal*, vol. 51, no. 6, pp. 1204–1222, 2008.

[109] H. Van Dijk, M. L. Van Engen, and D. Van Knippenberg, "Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance," *Organizational Behavior and Human Decision Processes*, vol. 119, no. 1, pp. 38–53, 2012.

[110] E. Kearney and D. Gebert, "Managing diversity and enhancing team outcomes: the promise of transformational leadership.," *Journal of applied psychology*, vol. 94, no. 1, p. 77, 2009.

[111] K. Bezrukova, S. M. Thatcher, and K. A. Jehn, "Group heterogeneity and faultlines: Comparing alignment and dispersion theories of group composition," *Conflict in organizational groups: New directions in theory and practice*, pp. 57–92, 2007.

[112] K. Bezrukova, S. Thatcher, K. A. Jehn, and C. S. Spell, "The effects of alignments: Examining group faultlines, organizational cultures, and performance.," *Journal of Applied Psychology*, vol. 97, no. 1, p. 77, 2012.

[113] D. A. Harrison, K. H. Price, J. H. Gavin, and A. T. Florey, "Time, teams, and task performance: Changing effects of surface-and deep-level diversity on group functioning," *Academy of management journal*, vol. 45, no. 5, pp. 1029–1045, 2002.

[114] P. Kollock, "The emergence of exchange structures: An experimental study of uncertainty, commitment, and trust," *American Journal of sociology*, vol. 100, no. 2, pp. 313–345, 1994.

[115] T. Yamagishi, K. S. Cook, and M. Watabe, "Uncertainty, trust, and commitment formation in the united states and japan," *American Journal of Sociology*, vol. 104, no. 1, pp. AJSv104p165–194, 1998.

[116] P. DiMaggio and H. Louch, "Socially embedded consumer transactions: for what kinds of purchases do people most often use networks?," *American sociological review*, pp. 619–637, 1998.

[117] D. H. Gruenfeld, E. A. Mannix, K. Y. Williams, and M. A. Neale, "Group composition and decision making: How member familiarity and information distribution affect process and performance," *Organizational behavior and human decision processes*, vol. 67, no. 1, pp. 1–15, 1996.

[118] S. L. Gaertner and J. F. Dovidio, *Reducing intergroup bias: The common ingroup identity model*. Psychology Press, 2014.

[119] A. Edmondson, "Psychological safety and learning behavior in work teams," *Administrative science quarterly*, vol. 44, no. 2, pp. 350–383, 1999.

[120] C. P. Earley and E. Mosakowski, "Creating hybrid team cultures: An empirical test of transnational team functioning," *Academy of Management journal*, vol. 43, no. 1, pp. 26–49, 2000.

[121] D. Lazer and N. Katz, "Building effective intra-organizational networks: The role of teams," 2003.

[122] D. J. Campbell and K. F. Gingrich, "The interactive effects of task complexity and participation on task performance: A field experiment," *Organizational behavior and human decision processes*, vol. 38, no. 2, pp. 162–180, 1986.

[123] H. Bui, V. S. Chau, M. Degl'Innocenti, L. Leone, and F. Vicentini, "The resilient organisation: A meta-analysis of the effect of communication on team diversity and team performance," *Applied Psychology*, vol. 68, no. 4, pp. 621–657, 2019.

[124] M. A. Campion, E. M. Papper, and G. J. Medsker, "Relations between work team characteristics and effectiveness: A replication and extension," *Personnel psychology*, vol. 49, no. 2, pp. 429–452, 1996.

[125] R. T. Keller, "Cross-functional project groups in research and new product development: Diversity, communications, job stress, and outcomes," *Academy of management journal*, vol. 44, no. 3, pp. 547–555, 2001.

[126] J. N. Cummings, "Work groups, structural diversity, and knowledge sharing in a global organization," *Management science*, vol. 50, no. 3, pp. 352–364, 2004.

[127] J. A. LePine, "Team adaptation and postchange performance: Effects of team composition in terms of members' cognitive ability and personality.," *Journal of applied psychology*, vol. 88, no. 1, p. 27, 2003.

[128] J. A. LePine, "Adaptation of teams in response to unforeseen change: effects of goal difficulty and team composition in terms of cognitive ability and goal orientation.," *Journal of Applied Psychology*, vol. 90, no. 6, p. 1153, 2005.

[129] J. R. Hollenbeck, A. P. Ellis, S. E. Humphrey, A. S. Garza, and D. R. Ilgen, "Asymmetry in structural adaptation: The differential impact of centralizing versus decentralizing team decision-making structures," *Organizational Behavior and Human Decision Processes*, vol. 114, no. 1, pp. 64–74, 2011.

[130] R. B. Davison, J. R. Hollenbeck, C. M. Barnes, D. J. Sleesman, and D. R. Ilgen, "Coordinated action in multiteam systems.," *Journal of Applied Psychology*, vol. 97, no. 4, p. 808, 2012.

[131] J. E. McGrath and T. W. Altermatt, "Observation and analysis of group interaction over time: Some methodological and strategic choicess," *Blackwell handbook of social psychology: Group processes*, pp. 525–556, 2001.

[132] A. M. Carton and J. N. Cummings, "A theory of subgroups in work teams," *Academy of management review*, vol. 37, no. 3, pp. 441–470, 2012.

[133] P. M. Blau, *Inequality and heterogeneity: A primitive theory of social structure*, vol. 7. Free Press New York, 1977.

[134] B. Meyer and A. Glenz, "Team faultline measures: A computational comparison and a new approach to multiple subgroups," *Organizational Research Methods*, vol. 16, no. 3, pp. 393–424, 2013.

[135] B. Meyer, A. Glenz, M. Antino, R. Rico, and V. González-Romá, "Faultlines and subgroups: A meta-review and measurement guide," *Small Group Research*, vol. 45, no. 6, pp. 633–670, 2014.

[136] P. D. Allison, *Event history and survival analysis: Regression for longitudinal event data*, vol. 46. SAGE publications, 2014.

[137] B. G. Kanki and H. C. Foushee, "Communication as group process mediator of aircrew performance.," *Aviation, Space, and Environmental Medicine*, 1989.

[138] P. S. Goodman and D. P. Leyden, "Familiarity and group productivity.," *Journal of applied psychology*, vol. 76, no. 4, p. 578, 1991.

[139] G. Littlepage, W. Robison, and K. Reddington, "Effects of task experience and group experience on group performance, member ability, and recognition of expertise," *Organizational behavior and human decision processes*, vol. 69, no. 2, pp. 133–147, 1997.

[140] P. J. Hinds, K. M. Carley, D. Krackhardt, and D. Wholey, "Choosing work group members: Balancing similarity, competence, and familiarity," *Organizational behavior and human decision processes*, vol. 81, no. 2, pp. 226–251, 2000.

[141] G. A. Okhuysen, "Structuring change: Familiarity and formal interventions in problem-solving groups," *Academy of Management Journal*, vol. 44, no. 4, pp. 794–808, 2001.

[142] R. Sosa and D. Albarran, "Supporting idea generation in design teams," in *DS 46: Proceedings of E&PDE 2008, the 10th International Conference on Engineering and Product Design Education, Barcelona, Spain, 04.-05.09. 2008*, 2008.

[143] B. Gokpinar, W. J. Hopp, and S. M. Iravani, "The impact of misalignment of organizational structure and product architecture on quality in complex product development," *Management science*, vol. 56, no. 3, pp. 468–484, 2010.

[144] R. Katz and T. J. Allen, "Investigating the not invented here (nih) syndrome: A look at the performance, tenure, and communication patterns of 50 r & d project groups," *R&d Management*, vol. 12, no. 1, pp. 7–20, 1982.

[145] D. Centola and M. Macy, "Complex contagions and the weakness of long ties," *American journal of Sociology*, vol. 113, no. 3, pp. 702–734, 2007.

[146] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," *Science*, vol. 337, no. 6092, pp. 337–341, 2012.

[147] S. Lehmann and Y.-Y. Ahn, *Complex spreading phenomena in social systems*. Springer, 2018.

[148] D. Guilbeault, J. Becker, and D. Centola, "Complex contagions: A decade in review," *Complex spreading phenomena in social systems*, pp. 3–25, 2018.

[149] N. Eagle and A. S. Pentland, "Reality mining: sensing complex social systems," *Personal and ubiquitous computing*, vol. 10, no. 4, pp. 255–268, 2006.

[150] N. A. Christakis and J. H. Fowler, *Connected: The surprising power of our social networks and how they shape our lives*. Little, Brown Spark, 2009.

[151] D. Easley, J. Kleinberg, *et al.*, *Networks, crowds, and markets*, vol. 8. Cambridge university press Cambridge, 2010.

[152] H. Fritz *et al.*, "The psychology of interpersonal relations," 1958.

[153] A. Rapoport, "Mathematical models of social interaction," 1963.

[154] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[155] S. P. Borgatti and M. G. Everett, "Models of core/periphery structures," *Social networks*, vol. 21, no. 4, pp. 375–395, 2000.

[156] P. Holme and J. Saramäki, "Temporal networks," *Physics reports*, vol. 519, no. 3, pp. 97–125, 2012.

[157] C. Ni, E. Smith, H. Yuan, V. Larivière, and C. R. Sugimoto, "The gendered nature of authorship," *Science Advances*, vol. 7, no. 36, p. eabe4639, 2021.

[158] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, "Networks beyond pairwise interactions: structure and dynamics," *Physics Reports*, vol. 874, pp. 1–92, 2020.

[159] D. M. Messick, "On the psychological exchange between leaders and followers," in *The psychology of leadership*, pp. 95–110, Psychology Press, 2004.

[160] E. P. Hollander, "Influence processes in leadership-followership: Inclusion and the idiosyncrasy credit model," *Advances in social & organizational psychology: A tribute to Ralph Rosnow*, pp. 293–312, 2006.

[161] L. Curral, P. Marques-Quinteiro, C. Gomes, and P. G. Lind, "Leadership as an emergent feature in social organizations: Insights from a laboratory simulation experiment," *PloS one*, vol. 11, no. 12, p. e0166697, 2016.