Game Changers: A Data-Driven Exploration of Team Design and Network Effects in Board Game Innovation

Computational Social Systems M.Sc.
Major in Social Data Science
Chair for Computational Social Sciences and Humanities
RWTH Aachen University

Master's Thesis

Handed in by:
Martin Koch, 413938

First examiner:
Prof. Jana Lasser

Second examiner:
Prof. Dr. Frank Piller

Registration date: 30.01.2024
Submission date: 29.07.2024

Zentrales Prüfungsamt/Central Examination Office

Eidesstattliche Versicherung
Statutory Declaration in Lieu of an Oath

___________________________ ___________________________

Name, Vorname/Last Name, First Name

Matrikelnummer (freiwillige Angabe)
Matriculation No. (optional)

Ich versichere hiermit an Eides Statt, dass ich die vorliegende Arbeit/Bachelorarbeit/ Masterarbeit* mit dem Titel

I hereby declare in lieu of an oath that I have completed the present paper/Bachelor thesis/Master thesis* entitled

____

____

____

selbstständig und ohne unzulässige fremde Hilfe (insbes. akademisches Ghostwriting) erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt. Für den Fall, dass die Arbeit zusätzlich auf einem Datenträger eingereicht wird, erkläre ich, dass die schriftliche und die elektronische Form vollständig übereinstimmen. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

independently and without illegitimate assistance from third parties (such as academic ghostwriters). I have used no other than the specified sources and aids. In case that the thesis is additionally submitted in an electronic format, I declare that the written and electronic versions are fully identical. The thesis has not been submitted to any examination body in this, or similar, form.

___________________________ ___________________________

Ort, Datum/City, Date

Unterschrift/Signature

*Nichtzutreffendes bitte streichen

*Please delete as appropriate

Belehrung:
Official Notification:

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

Para. 156 StGB (German Criminal Code): False Statutory Declarations

Whoever before a public authority competent to administer statutory declarations falsely makes such a declaration or falsely testifies while referring to such a declaration shall be liable to imprisonment not exceeding three years or a fine.

§ 161 StGB: Fahrlässiger Falscheid; fahrlässige falsche Versicherung an Eides Statt

(1) Wenn eine der in den §§ 154 bis 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

(2) Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Abs. 2 und 3 gelten entsprechend.

Para. 161 StGB (German Criminal Code): False Statutory Declarations Due to Negligence

(1) If a person commits one of the offences listed in sections 154 through 156 negligently the penalty shall be imprisonment not exceeding one year or a fine.

(2) The offender shall be exempt from liability if he or she corrects their false testimony in time. The provisions of section 158 (2) and (3) shall apply accordingly.

Die vorstehende Belehrung habe ich zur Kenntnis genommen:

I have read and understood the above official notification:

___________________________ ___________________________

Ort, Datum/City, Date

Unterschrift/Signature

Contents

Abstract ix

Überblick xi

1 Introduction 1

2 Background 5

2.1 On Measuring Creativity 5

2.1.1 Novelty and Usefulness 5

2.1.2 Innovation and Creativity 6

2.1.3 Operationalizing Novelty and Usefulness 6

2.1.4 Integrating and Theorizing Conceptualizations 6

2.2 Team Effectiveness and Collaborative Creativity 8

2.2.1 Team Effectiveness Framework 9

2.2.2 Model of Collaborative Creativity 11

2.3 Networks (Dynamics) and Team Creativity 14

2.3.1 Consequences of Network Dynamics 15

2.4 Research Gaps 16
2.5 Research Context 17
2.6 Related Work 19
2.7 Research Goals 22
2.7.1 Research Questions 22
2.7.2 Research Hypotheses 22

3 Methods **25**

3.1 BoardGameGeek Dataset 25
3.2 Data-Driven Analysis of Teams 28
3.2.1 From Scientific Research to Board Game Innovation 29
3.2.2 BGG Data Traces–Analytical Potential 29
3.2.3 Why Do We Need Statistical Network Models? 31
3.3 Model Framework 33
3.3.1 Background 34
3.3.2 Relational Hyper Event Model (RHEM) 35
3.3.3 Relational Hyper Event Outcome Model (RHOM) 39
3.3.4 Hyperedge Statistics for RHEM and RHOM 40
3.4 Model Specification 46
3.5 Data Cleaning and Preparation 50

4 Results **55**

4.1 Overview 55

4.2 Hyperedge and Descriptive Statistics Analysis 56
4.3 Hyperedge Statistics and Collaboration Probability (H1) 60
4.4 Correlation between Collaboration Probability and Creative Outcomes (H2) . . . 64
4.5 Comparison between High-Creativity and Low-Creativity Teams (H3) 68
4.6 Summary of Findings 71

5 Discussion **73**

6 Conclusion **79**

A Code Repository **83**

B Supplementary Information **85**

Bibliography **93**

List of Figures and Tables

3.1 Summary of BGG Dataset 25
3.2 Evolution of the Number of Designers, Mechanics, and Mechanics Combinations 26
3.3 Prevalence of Different Team Sizes Over Time 27
3.4 Similarity of Collaborations in Scientific Research and Board Game Innovation . 30

4.1 Cleaned BGG Dataset 56
4.2 CoxPH Coefficients 62
4.3 Correlation between Collaboration Probability and Creative Outcomes 66
4.4 Comparison between High-Creativity and Low-Creativity Teams 69

B.1 VIF Scores for OLS on Novelty. 88
B.2 VIF Scores for OLS on Usefulness. 89
B.3 VIF Scores for CoxPH on Novelty. 90
B.4 VIF Scores for CoxPH on Usefulness. 91

Abstract

This master's thesis investigates the dynamics of team formation and creative success within the board game industry, utilizing data from BoardGameGeek. By applying the novel Relational Hyper Event Model (RHEM) and Relational Hyper Event Outcome Model (RHOM), we identify key factors influencing collaboration and their impact on creative outcomes. The analysis reveals that prior collaborations and tenure similarity significantly drive team formation, while prior success and diverse mechanical experience correlate with higher creative performance. High-creativity teams are larger, exhibit more skill-related diversity, and have greater mechanical experience. These findings underscore the importance of balancing familiarity and diversity to foster innovation. Methodological insights and future research directions are discussed, highlighting the potential of RHEM and RHOM in practical research contexts.

Überblick

Diese Masterarbeit untersucht Netzwerkeffekte in der Teambildung und deren Zusammenhang mit kreativem Erfolg in der Brettspielbranche unter Nutzung von Daten von BoardGameGeek. Durch die Anwendung des kürzlich entwickelten Relational Hyper Event Model (RHEM) und des Relational Hyper Event Outcome Model (RHOM) werden Faktoren identifiziert, welche die Zusammenarbeit und ihre Auswirkungen auf kreative Ergebnisse beeinflussen. Die Analyse zeigt, dass frühere Kooperationen und ähnliche Erfahrungszeiträume unter Brettspieleentwicklern häufiger zur Teamformation führen, während vorheriger Erfolg und Erfahrungen in der Entwicklung von vielseitigen Spielen mit höherer kreativer Leistung korrelieren. Hochkreative Teams sind größer, weisen mehr job-relevante Vielfalt auf und haben größere Spieleportfolios. Diese Erkenntnisse unterstreichen die Bedeutung der Balance zwischen Vertrautheit und Veränderung zur Förderung von Kreativität. Methodologische Erkenntnisse und zukünftige Forschungsperspektiven werden diskutiert, wobei das Potenzial von RHEM und RHOM in praktischen Forschungskontexten hervorgehoben wird.

Chapter 1

Introduction

In today's world, complex, non-routine tasks increasingly require team collaboration rather than individual effort. This is particularly true in creative industries such as board games, where generating and implementing innovative solutions demands significant resources.

Research on creativity and innovation spans diverse fields, including psychology, sociology, and network science. Organizational studies have examined how individual capabilities and team processes influence creativity, while social network theory has explored the impact of social interactions and embeddedness on creative outcomes.

Team creativity, defined by the novelty and usefulness of its products, requires balancing divergent thinking for generating ideas and convergent thinking for refining them. Team characteristics such as diversity and experience can facilitate or hinder this process. Diverse teams offer a range of perspectives but may face coordination challenges, while integrating newcomers can introduce fresh ideas but also risks underperformance.

Most research in corporate and scientific domains assumes that team formation is directed by superiors. However, in creative industries, team formation is often self-selected. This difference necessitates a unique approach to studying team dynamics in these contexts. Understanding how teams form and achieve creative success in such environments is crucial.

This thesis investigates the factors influencing team formation and creative performance in the board game industry. By integrating previous research with data-driven methods, we explore the characteristics of successful teams and the dynamics of their collaborations. Methodologically, we employ a combination of network analysis and statistical modeling to capture the complexity of team interactions and their outcomes. This approach allows us to account for both the relational and temporal aspects of the formation and performance of creative teams.

We aim to identify key factors that foster effective teamwork and drive creativity in board game design, contributing to a broader understanding of innovation in self-organized teams.

The remainder of this thesis is structured as follows:

- **Chapter 2: Background** - This chapter provides a comprehensive review of existing research on team creativity and innovation. It covers theories and findings from psychology, sociology, and network science, highlighting their relevance to the board game industry.
- **Chapter 3: Methods** - This chapter describes the dataset used for the study, sourced from BoardGameGeek. It details the data cleaning process, the extraction of relevant features, and the statistical methods employed to analyze team formation and creative outcomes.
- **Chapter 4: Results** - This chapter presents the findings of the study. It discusses the main drivers of team formation, the relationship between these drivers and creative success, and the differences between high-creativity and low-creativity teams.
- **Chapter 5: Discussion** - This chapter interprets the results in the context of existing literature. It explores the implications of the findings for understanding team creativity in the context of the board game industry.
- **Chapter 6: Conclusion** - This chapter summarizes the key contributions of the thesis, discusses its limitations, and suggests directions for future research.

Chapter 2

Background

2.1 On Measuring Creativity

Creativity research in the organizational context focuses on the "generation of new and useful ideas" [Amabile, 1983] to achieve and maintain success. An idea, in this sense, as a performance outcome, can take the form of a product, service, or process and is distinct from regular performance in that it does not merely execute predefined procedures but also involves new and better ways of achieving a certain objective.

2.1.1 Novelty and Usefulness

The creative process of generating and evaluating ideas encompasses two core constructs: novelty and usefulness. Since [Amabile, 1983], these constructs have been widely accepted in creativity research [Zhou and Hoever, 2023; Byron et al., 2022; Pollok et al., 2021].

Novelty Novelty involves fundamentally new ways of solving a problem, such as new combinations of essential parts of previous ideas. It differs from iterative improvements, which are incremental changes rather than fundamental shifts.

Usefulness A new idea or product must also be useful. Usefulness is judged by the community that uses and evaluates the ideas. As it is socially embedded, perceptions of usefulness are context-dependent and can change over time.

2.1.2 Innovation and Creativity

Innovation and creativity should be conceived as different stages of the same process: creativity as the generation of new ideas and innovation as their successful implementation. While this suggests a linear progression (first the idea, then the implementation), some research increasingly views this process as iterative or chaotic. This complexity makes it difficult to measure, leading researchers to often use both constructs equivalently despite their theoretical differences [Paulus et al., 2011]. In this work, we acknowledge that creativity is a process. However, we use innovation and creativity equivalently due to the practical difficulty in distinguishing them and their irrelevance to many hypotheses.

2.1.3 Operationalizing Novelty and Usefulness

Creativity research studies the antecedents of novelty and usefulness, which are inherently bound by time, place, and context [Amabile and Pratt, 2016].

Novelty Novelty is often operationalized in terms of the number of ideas generated, the number of categories a set of ideas can be applied to, or the statistical rarity of an idea compared to other ideas or products. This can include uniqueness (specific to a domain), unexpectedness (e.g., ideas imported from another domain), or non-typicality (relative to a given domain).

Usefulness Usefulness can be measured in terms of feasibility, value to an audience, or sufficiency in addressing a proposed problem within a domain.

Despite the consensus on the two dimensions of creativity, as a recent paper by [Harvey and Berry, 2023] highlights, the field has seen "substantial diversity in how novelty and usefulness are defined or operationalized." This diversity leads to disagreements that, while not due to construct incompatibility, are difficult to untangle from contradictory findings in the already complex theories [Zhou and Hoever, 2023]. These disagreements can be explained in terms of different "forms" of creativity.

2.1.4 Integrating and Theorizing Conceptualizations

According to [Harvey and Berry, 2023]'s distal framework, perspectives on creativity differ in theoretical roots, sources of creativity, the relationship between novelty and usefulness, and the creative process.

The **maximization perspective** suggests that creativity is highest when novelty and usefulness are independently maximized. Rooted in evolutionary theory and stage models, this view treats novelty and usefulness as separate stages in a linear process.

For example, a board game designer might first brainstorm highly novel game mechanics[1] and then refine the rules to ensure the game is playable and enjoyable.

The **balance perspective** posits that creativity peaks when novelty and usefulness are balanced at moderate levels. Based on conflict and paradox theories, this perspective sees novelty and usefulness as inversely related, requiring a balance for optimal creativity.

For instance, a designer might develop a novel mechanic balanced with familiar ones, adjusting the game based on feedback to ensure it is both innovative and playable.

The **integration perspective** argues that creativity is maximized when novelty and usefulness are interdependent and positively related. This view, grounded in synthesis and brokerage theories, sees creativity as an integrative process where novelty and usefulness enhance each other.

For example, a designer might evolve the game's mechanics during the design process, integrating continuous feedback to ensure novelty and playability enhance each other.

Although the creative process varies, we can only observe the final product, using metrics for both novelty and usefulness. Thus, products from different perspectives will reflect distinct balances of these criteria: highly novel and refined (maximization), moderately innovative and balanced (balance), or seamlessly intertwined novelty and usefulness (integration). The specific quantitative measures for assessing these constructs will be detailed in the methods section.

We note that what it means for a product to be creative can, among other things, depend on the form of creativity. Understanding which form of creativity is at play affects the conceptualization and subsequent measurements. Even though a single form hardly ever captures creativity completely, in order to make research more coherent, considering the following questions during operation helps to theoretically motivate the operationalization of a specific form:

- What are the supposed drivers of creativity?
- How do novelty and usefulness relate (independent, paradox, interdependent)?
- How far away are creators from the evaluative context?
- How is the creative process perceived (linear, iteration, integration)?

[1] Fundamental and characterizing building block of a game.

2.2 Team Effectiveness and Collaborative Creativity

This section will establish a conceptual baseline with respect to the study of teams and teamwork and further introduce a framework that relates team design with creativity. In order to do this, we will heavily draw from [Lantz Friedrich and Ulber, 2017]'s approach as a guiding framework. Despite the fact that they purport a model (or models) that is mostly used in psychology-based research, it generalises well to related disciplines, at least as an organising framework and to establish common constructs.

From Groups to Teams

Teams as a way of organising work is fundamentally motivated by the idea of synergy – i.e., 'the whole is more than the sum of its parts.' This implies that the social interactions that take place during teamwork not only encompass outcomes which directly meet the immediate task requirements, but they also attain outcomes for individuals, teams, and organisations. For example, "fulfilling individuals' needs for growth and health [can] be a source of motivation and satisfaction" which [Lantz Friedrich and Ulber, 2017] in turn relates with organisational level outcomes such as productivity, efficiency, and innovation. Unfortunately, there are at least as many challenges as opportunities in the social processes that could inhibit or even prevent success. Examples include team members not getting along or a lack of feeling responsible for a specific outcome (for example, because blame for a potential failure is directed at the group and not the individual). While there is significant overlap in the study of group phenomena relevant to all of social psychology, an essential distinction organisational research makes is between a group and a team:

- A team is not merely a group of people (e.g., employees working in the same space), but requires members to interdepend with respect to a work task and to engage in social interaction. More completely, research (Lantz Friedrich and Ulber [2017]; Levi [2001]) expects teams to consist of members who:
  - know they are part of the team (boundedness),
  - depend in some way on other members to complete their task (task interdependence),
  - jointly define goals (shared objectives), and
  - reflect on performance (reflexivity).

As this definition of teams consists of interpersonal and task dimensions, organisational research on teams studies how interpersonal and task-work processes interrelate. One prominent line of team research focuses on team effectiveness. Noting that creative performance is more complex than routine performance as later discussed, the core ideas are the same and this lens commonly serves as a foundation for creativity research.

2.2.1 Team Effectiveness Framework

The input-process-output (IPO) model, developed in 1964, has since served as a foundational framework for comprehending how team dynamics within organizations affect their performance (Ilgen et al. [2005]). It involves three key components (which also name the model):

- inputs, representing individual and group attributes, the task and context characteristics,
- processes, encompassing coordination, communication, team interaction, and
- outputs, comprising measurable group performance and subjective member reactions.

The IPO model understands team performance as a "static linear process" (Lantz Friedrich et al. [2020]) that could be interpreted causally, i.e., inputs affect processes, processes affect outcomes. As a model from organisational psychology, team processes are analysed using already established constructs: social processes, cognitive processes, and motivational processes.

While this model appears simplistic, the simultaneous occurrence of team processes within the IPO model can become very complicated. Yet, by assuming static relationships between the model's elements, it fails to account for the complex nature of team dynamics, particularly how they develop over time and through interaction. To account for this, newer models (like the IMOI) introduced feedback loops between inputs, processes, and outputs (Ilgen et al. [2005]). Furthermore, inputs, processes, and outputs are no longer assumed to be causally related but, in line with increasing evidence, the model proposes that inputs are mediated by processes in the achievement of some output, hence the new name Input-Mediator-Output-Input (IMOI).

The Input-Mediator-Output-Input (IMOI) model emerged as a new foundational framework, offering a more comprehensive understanding of the intricate interplay between inputs, processes, and outcomes within teams. Looking at team effectiveness through a multi-level lens, the input, mediators, and outputs describe (team) characteristics at various levels which also interrelate through feedback loops.

Specific levels of the input and output describe the individual, the team, or the organisational characteristics. As can be seen in the figure, individuals (the micro-level) are embedded in a team (the meso-level), the team is embedded in an organizational context (the macro-level), and contexts on different levels influence each other along the developmental process (Ilgen et al. [2005]).

Input Levels

- **Macro**: Organizational support, leadership, and HR practices.
- **Meso**: Interdependence, task design, team size, and composition.
- **Micro**: Individual KSAOs (Knowledge, Skills, Abilities, and Other characteristics).

Meso-level input factors such as team composition, diversity, and size are the main focus of this work and will be the topic of the next section. Similarly to the inputs, the output can be considered at different levels.

Outcome Levels

- **Macro**: Results tied to organizational goals, such as innovation.
- **Meso**: Team viability and learning outcomes.
- **Micro**: Addressing individuals' needs for job satisfaction, health, and personal growth.

Notwithstanding conceiving of inputs as prerequisites, e.g., having the necessary skills for a task, superficially similar teams can perform differently, which the IMOI model tries to explain by mediator variables. Given the prerequisites for fulfilling a task, mediators describe how emergent processes such as team or leadership processes affect task outcomes. Routinised patterns of interaction can affect the way inputs are turned into output and further self-reflection on joint experiences informs future interactions. These processes are inherently difficult to capture but have been found to largely determine team effectiveness. Key team processes encompass cognitive, emotional, and behavioral dimensions (Ilgen et al. [2005]).

Mediators

- Cognitive processes involve building shared meaning through team learning and reflexivity, leading to emergent states of Team Mental Models (TMM) and team climate.
- Affective processes address emotions, cultivate trust, cohesion, and team efficacy, and manage conflicts that may arise within the team.
- Behavioural processes focus on achieving goals and resolving task demands through coordination, cooperation, and effective communication, determining team performance and outcomes.

While IMOI is generally accepted as the more appropriate model for team effectiveness, IPO-based models are still being deployed for their simplicity in measurement. The advantages IMOI brings are particularly relevant when studying more complex performance such as creativity performance.

2.2.2 Model of Collaborative Creativity

The study of creativity at the team level has only gained traction in the last decade. Early seminal works Amabile [1983] on organizational creativity assumed that (small) team creativity functioned equivalently to individual creativity. Until 2010, the primary focus was not on the team itself but on the factors within the team that inhibited or benefited individual creativity. Recently, this has changed, and team creativity research has seen rapid growth. One of the earlier models is the collaborative creativity framework by Paulus et al. [2011]. This model "considers how many different factors and variables influence the cognitive, social, and motivational processes that underlie collaborative creativity." Byron et al. [2023]

The model is largely inspired by earlier work on team effectiveness but recognizes that "creative" performance (as described in more detail in measuring creativity) has specific challenges compared to "routine" performance. First, the definition of creativity and innovation is a social criterion and, as such, is bound by time, context, and place, involving more uncertainty than mere benchmarking. Traditional benchmarking typically involves comparing performance metrics or outcomes against predefined, often standardized metrics. Second, inherent to creativity is a paradoxical tension between the idea generation and idea implementation phases, both of which are part of a "creative product." This tension is sometimes also framed in terms of exploration and exploitation Harvey and Berry [2023], which challenges a work group more than a regular task.

Team creativity, as defined by the model, involves the collaborative generation of new and feasible ideas by a group of individuals. This implies that creative solutions necessitate team members working together. Understanding team design is important, as it establishes the basis for team processes that eventually produce outcomes. Collaboration between team members is explained through social, cognitive, and motivational processes (equivalent to the IMOI model).

To achieve highly creative outcomes, team members are required to:

- Engage in social processes such as freely sharing ideas, exchanging information, discussing diverse viewpoints, and managing conflict.
- Participate in cognitive processes by broadly searching for and attending to novel ideas, and then combining and elaborating on these ideas.

- Involve motivational processes to maintain sufficient motivation—both intrinsically and extrinsically—and prevent potential group motivational losses, such as social loafing.

In addition to various processes, the model integrates decades of research on the antecedents of team creativity and comprises a comprehensive list of over 30 input variables. Despite the high number of factors, the main complexity of the model lies in how different design aspects relate to creativity. Each aspect of team design is related to a specific set of social, cognitive, and motivational processes. For example, with team size: More members imply more skills and expertise, which is beneficial to ideation, but communication within a group becomes less efficient.

Furthermore, most factors of team design have some limiting conditions. For instance, while team cohesion can enhance trust and collaboration, excessive cohesion may lead to conformity and reduce creative dissent. Similarly, diversity within a team broadens perspectives and fosters innovation, yet too much diversity can introduce communication barriers and conflicts that hinder group dynamics. These and other factors require careful consideration and balance to optimize team creativity effectively. Most research on team creativity adopts a "creativity as balance" account, in which creative products balance novelty and usefulness.

Meta-Analysis: Team Design and Team Creativity

Expanding upon this, a recent meta-analysis (Byron et al. [2023]) underscored the necessity of simultaneously considering the collective interaction processes that underlie collective creative outcomes. This emphasises that to understand the intricacies of collective creativity, various factors need to be integrated. These factors can broadly be grouped into team composition, task structure, and organisational support – all of which they subsumed under the umbrella term "team design".

While task structure and organizational support are important components of team design, they are not our primary focus due to their stability in our later case study context. Nevertheless, as they are integral parts of team design we will give a brief overview:

- Task Structure involves the organization, allocation, and coordination of tasks within the team, aligning individual and collective goals to facilitate effective collaboration.

- Organizational Support covers the resources provided, support from leadership, and the autonomy granted to the team, along with the operational and environmental conditions that affect team functioning. This support is crucial for enhancing productivity and encouraging innovation.

As previously stated, understanding the intricate relationship between team design and its impact on team creativity and innovation is crucial because team design establishes the basis for team processes.

Here's a summary of the key points, hypotheses, and findings related to Team Composition:

- **Team Tenure**: Team tenure refers to the duration members have been together in a team. The relationship between team tenure and creativity and innovation is complex and potentially curvilinear. While longer-tenured teams generally benefit from enhanced trust, psychological safety, and reduced conflict, boosting creativity and innovation, there might be a curvilinear effect where teams of moderate tenure experience lower cohesion and increased conflict before eventually stabilizing and improving as they gain familiarity. Further, longer-tenured teams can develop efficient transactive memory systems, but they might also suffer from groupthink or a resistance to new information.
- **Team Size**: The model predicts a nuanced relationship between team size and creativity, which could be linear or curvilinear. Larger teams might offer a wealth of perspectives and cognitive resources, enhancing creative outputs. However, they also face challenges such as coordination difficulties, diminished personal accountability, and potential for free-riding, which could reduce creativity and innovation.
- **Team Diversity**: The impact of diversity on team creativity and innovation varies based on the type of diversity. Job-related diversity, like functional or educational background diversity, typically aids creativity by bringing varied knowledge and perspectives. In contrast, demographic diversity might hinder team processes due to increased conflict and impaired relationships among members from different demographic backgrounds, although under certain conditions, it could enhance creativity.

The meta-analysis on 134 field studies and 35 student studies between 2010 and 2020 examined the role of team design. With respect to the consequences of team composition on team creativity and innovation they:

- find a curvilinear relationship between team tenure and creativity, suggesting that while newly formed teams and long-established teams exhibit higher creativity and innovation, those with moderate tenure may struggle with cohesion and conflict.
- Regarding team size, the analysis indicates no significant overall impact on creativity and innovation, pointing to a balance of positive and negative effects from increased cognitive resources and coordination challenges.
- For job-related diversity, the findings consistently show a positive influence on creativity and innovation due to the diverse skills and perspectives it brings.

- However, demographic diversity shows mixed effects, with generally weak or null relationships with creativity and innovation, suggesting that its benefits may be context-dependent.

Overall, the literature highlights that the elements of team design—tenure, size, and diversity—have complex and often nonlinear relationships with team creativity and innovation. These dynamics are crucial for understanding how teams can be structured and managed to maximize creative outcomes. There have been other calls Zhou and Hoever [2023] to further explore these relationships in more depth, particularly the conditions under which demographic diversity can either hinder or enhance team performance or to explore the complexities of team tenure.

2.3 Networks (Dynamics) and Team Creativity

Social network studies within the context of team creativity focus on describing and explaining how the structure of relationships among team members influences creative outcomes. Traditional research has largely examined static networks [Perry-Smith and Mannucci, 2017; Perry-Smith and Shalley, 2003], exploring phenomena such as the small worldliness of creative networks [Baten, 2021], brokerage positions [Juhász et al., 2020], and access to information [Soda et al., 2021; Mannucci and Perry-Smith, 2022]. These studies assumed that network structures express stable relational states. However, recent shifts in the field acknowledge the dynamic nature of social networks, which change over time and through interactions. This shift introduces the concept of "network dynamics," encompassing the formation, reproduction, and transformation of ties. A critical challenge in this stream is translating network theories, which typically assume static states, into frameworks that account for relational events and their effects on creativity [Chen et al., 2022; Fleming et al., 2007; Klonek et al., 2019; Rosing et al., 2018].

For example, social network theory traditionally examines the structure of relationships within teams, emphasizing the diversity of ties for fostering creativity. However, when considering network dynamics, it becomes essential to understand how these ties change over time and through specific interactions. Granovetter's strength of weak ties theory, which highlights the role of weak and bridging ties in introducing novel information, must be adapted to account for the timing and sequence of these interactions. Similarly, Burt's structural hole theory and network embeddedness theory examine how bridging gaps in networks and the overall pattern of connections impact creativity. Yet, these theories must also consider how dynamic interactions, rather than static positions, influence creative processes.

Research in network dynamics within team creativity faces the complexity of co-creating network states and events, where roles and behaviors are intertwined. This complexity is further compounded by the need to understand how changes in the composition of the network, driven by contextual factors, actor attributes, and relational factors, affect team performance and creative output. Contextual factors include organizational contexts and physical workspaces that shape relationship opportunities over time. Actor attributes focus on demographic or task-related homophily, and relational factors consider local tie dependencies, such as the tendency for triadic closure.

2.3.1 Consequences of Network Dynamics

A recent meta-review on network dynamics identified three distinct subcategories in the emerging network literature, which are best understood in terms of what the networks are assumed to describe. Most commonly, network data is posited to capture relationships among social actors. It follows that longitudinal network data captures network change (1) or how the respective relationships change, i.e., the "formation, reproduction, and transformation" of ties Chen et al. [2022]. When interpreting networks as relationships, we presume relational states (e.g., friendship ties), which is often unwarranted, considering the nature of digital trace data. This sort of data is a result of social (inter)actions, which should rather be conceived as relational events (2). In contrast with states, when studying relational events, one tries to understand the sequencing and timing of actions. While many studies simply aggregate (e.g., count) the amount of interactions to infer relationships for a specific time period, the distinctions of network states and events are complex. From a more theoretical perspective, it can be said that they co-create each other. Simply put, one could conceptualize network states as roles and network events as acts or behaviors. The third conceptualization is in terms of how nodes and network co-evolve (3).

Compared to other organizational research streams that pertain to a (relatively) established set of psychological/social processes (covered in the previous section), theories that try to explain network dynamics are less unified. Chen et al. [2022] identified three drivers of network dynamics, each either at interpersonal or interorganizational levels. Contextual factors include memberships and roles/positions in processes, whether actors live in the same region, and other boundary-setting organizational contexts which influence ties change. An example could be how patterns of career mobility shape interpersonal network dynamics, or how the physical workspace shapes opportunities to foster relationships over time. Theories on actor attributes can ask for the consequences of demographic or task-related homophily within a team Ertug et al. [2022]. Theories reflect the impact of preferences or highlight competitive advantages in terms of network position in the context of social capital, risk, status, and related constructs. Theories on relational factors theorize upon local dependencies of ties. The classic example is the question of whether "we tend to befriend our friends' friends"; it becomes more likely that two actors who share a common friend eventually befriend each other as well, sometimes referred to as triadic closure. More generally, it is concerned with how current network dynamics/constellations affect subsequent network dynamics and include works on structural holes and balance theory. It is important to note that these drivers can be highly interrelated, and this separation is rather conceptual.

The block of outcomes relates network dynamics with attitudes, actions, and performances of network actors. For the scope of this work, we will focus on outcomes as performance, which in their framework subsumes research on innovation and creativity. One controversial topic is concerned with endogeneity, i.e., the question of whether network dynamics really precede performance. For example, even though it is often theorized that brokering positions improve performance, more recent work suggested that high performance ascertains future brokering positions, highlighting the complex nature. Nevertheless, theories on the consequences of network dynamics highlight processes that are difficult to encode in alternative frameworks. This includes how the age of ties possibly encodes historic conditions, how oscillating patterns of specific network changes relate with innovation, or how the benefits of certain network positions depend on the stability of the network as a whole. Important to this work is the question of how changes in the composition of the network affect team performance. For example, social resource theory posits that resource availability leading to innovation is more likely to be powered by changes in composition rather than some beneficial but fixed structure. They argue that in contexts of fast-paced innovation, to uphold task-relevant diversity within a team, the right amount of compositional changes is required. This could, for example, be explored by moving the interpretative framework from network states to events.

Overall, Chen et al. [2022] concludes there is a lack of research investigating the relationship of network dynamics with outcomes. Importantly, the mechanisms that explain relational events might be studied using classical network mechanisms of resource access, trust, power, and signaling but need to be adapted. In section 3, we will focus on statistical regularities of variables that relate to organizational psychology and key elements of network theories, such as access, trust, power, and signaling.

2.4 Research Gaps

Teamwork today is pervasive and has been studied from various perspectives. A recurring question is what makes some teams achieve better creative outcomes than others. Previous chapters depict how research on team creativity provides multidisciplinary insights into designing and supporting teams, primarily drawing from social psychology and network theory. However, much of this research was conducted within the organizational settings of larger firms, often implicating a specific sort of team that may not translate well to other contexts.

While the criteria for defining a team—boundedness, task interdependence, and reflexivity—are broad enough to encompass multiple forms of teams, membership is often assumed to be relatively stable. This captures team selection and formation within a specific type of team, such as a stable group of 2-10 individuals working on one project at a time within a company.

Some research considers other team contexts like ad-hoc teams (spontaneously formed), emergency response teams (newly formed and high-pressure), and fluid teams (rapidly composed without familiarity and part of many teams at once) [Linhardt and Salas, 2023].

In the context of board game developers and related creative industries, one distinctive characteristic is that members self-select whom to work with on a project-to-project basis, often resulting in frequent partial or complete membership changes that reflect joint experiences. Accounting for self-selection and frequent, complex membership changes as collective interaction processes that underlie collective creative outcomes is hardly addressed in prior research on creativity. By investigating how patterns of collaboration affect creative outcomes, we respond to Chen et al. [2022]'s call to study the complexities of team tenure and Lerner and Lomi [2023]'s criticism that temporal and higher-order effects in social networks should be used "as an opportunity to develop and test additional theories about the structure and dynamics of social interaction."

Next, we will introduce the research context (the board game industry), review related work, and develop our hypotheses for this study.

2.5 Research Context

"Like novels, games don't come out of the blue. They come from a given designer, at a given time, in a given social situation." —Bruno Faidutti as quoted in Woods [2012]

In our case study, we will apply insights from creativity research to board game innovation to better understand the social situations that give rise to creative outcomes. Specifically, we will explore a dataset on board game designers (detailed in the methods section) to see how teams form around the production of games and how this formation affects the creative outcome.

Board Game Industry and History

Since the early 1990s, the board game industry has experienced rapid growth, driven by both technological and cultural developments. Technologically, the internet enabled information-sharing among designers and consumers without the need for co-location. Culturally, developments particularly in Germany, such as the emergence of related magazines, conventions, and awards (e.g., "Spiel des Jahres"), have fostered idea exchange and community growth. By 2018, the industry's global revenues reached approximately $12 billion, reflecting its booming success, which already more than 20 years ago was thought to have reached its "golden era."

Before discussing recent research on board game innovation, we will briefly introduce the research context. Throughout history, board games have been integral to most human societies. These games consist of various elements, rules, and themes that players interact with to achieve specific goals. Narrating the history of board games often involves exploring their development and innovation, describing how popular game designs evolve in terms of characteristics, goals, and player interactions. A comprehensive review of board game culture is beyond the scope of this work. Instead, we will highlight aspects relevant to our study. Classic (old to ancient) games, often not attributable to a specific designer, are characterized by strict rules, high complexity, and educational purposes. Examples such as chess and go remain popular today. Post-WWII, several streams of board game designs emerged commercially. One stream, particularly popular in the US and UK, focused on thematic elements with less complex gameplay, exemplified by war-games. Another stream included simple party games with minimal player interaction, like Monopoly.

In the 1990s, a design stream known as "Eurogames" gained global popularity and is hitherto pervasive. Driven by cultural and technological developments in Europe, these games emphasized player interaction and entertainment. Eurogames are built around mechanics that facilitate social play, characterized by accessible themes, simple rules, constrained playing times, and non-confrontational interactions [Woods, 2012; Lutter and Weidner, 2021; Pollok et al., 2021].

Mechanics constitute the primary elements that control player movements and their interactions during the course of the game. Defined as the "various actions, behaviors, and control mechanisms afforded to the player within a game context," these mechanics "combine to form a system through which players interact, make choices, and experience the game" [Sousa et al., 2021; Hunicke et al., 2004]. One notable example is the trading mechanic in "Settlers of Catan," possibly the most famous Eurogame to date. While themes help contextualize and remember game rules, the primary enjoyment comes from player interactions facilitated by game mechanics. For this work, we will focus on board games as "creative" combinations of mechanics, largely neglecting the less tangible aspect of themes. However, the importance of game themes remains a topic of debate among designers.

Board Game Industry and Collaborative Innovation

In the board game industry, mechanics are creatively combined by designers during the development process. As in other creative industries, success in this highly dynamic and rapidly growing field relies on knowledge and creativity rather than substantial capital investments. While board game designing does not require a specific technical background, certain knowledge and skills are essential. Understanding game mechanics, player psychology, and thematic integration are critical competencies that directly influence the quality of the final product Pollok et al. [2021]. The initial creative idea must be effectively translated into an engaging and playable game, and this is closely reflected in the end product that players evaluate.

Even though board games are still mostly created by single developers, they are increasingly often the result of team efforts. According to some designers [Cortinovis and van der Wouden, 2021], working in a team increases effectiveness and leads to better results. Yet, these teams are much smaller compared to those in video game or film production, probably due to less coordination overhead in both development and distribution. In the case of teamwork, designers in the board game industry are not assigned to collaborate on projects. Instead, they have significant agency to identify and select their team members. This autonomy allows for the formation of teams based on mutual interests, complementary skills, and shared creative visions [Cortinovis and van der Wouden, 2021]. If and when self-selected collaborations lead to more creative outcomes will be the topic of this work.

Overall, the board game industry exemplifies a domain where small, agile teams can produce highly innovative and successful products. The industry's growth and dynamism are fueled by the creativity and knowledge of individual designers and their ability to collaborate effectively. As the industry continues to evolve, the emphasis on creativity and designer agency will likely remain a cornerstone of board game innovation.

2.6 Related Work

The board game industry has lately attracted attention from creativity researchers due to its economic impact and the growing online community of designers and players. BoardGameGeek (BGG), a community-run database and forum, has provided a rich source of observational data on creative collaborations. This data includes designers' publications, collaborations, user ratings, and established game categorizations, offering a structured way to compare games. This section reviews four studies leveraging BGG data to explore creative collaborations, examining their research questions, underlying theories, operationalization of creativity, and findings.

[Wachs and Vedres, 2021]

In their study "Does Crowdfunding Really Foster Innovation? Evidence from the Board Game Industry," the authors explore whether crowdfunded board games are more innovative than traditionally funded ones. The study is based on theories of gatekeeping and institutional innovation. Innovation is defined as novelty (novelty and distinctiveness) and prototypicality (resonance) in terms of game mechanics, with distinctiveness measured by the average hamming distance to previous games, novelty by the minimum hamming distance, and resonance by the difference between past and future distinctiveness. The researchers control for selection bias into crowdfunding by considering team (size, number of newcomers) and game (complexity, year, genre) covariates. Their findings suggest that crowdfunding supports innovation by enhancing novelty and distinctiveness and is a predictor of future trends in the industry.

[Pollok et al., 2021]

They investigate the reasons user teams might outperform professional teams in translating diversity into creativity in their study titled "Knowledge diversity and team creativity: How hobbyists beat professional designers in creating novel board games." This study is grounded in Information Processing Theory and the concept of Representational Gaps. Creativity is measured separately for novelty (distinctiveness of game mechanics combination) and usefulness (ratings). The researchers regress diversity on each creativity component and explore how team type (user vs. professional) moderates these relationships. They control for exogenous team (status, size, gender) and game (complexity, version, language) covariates. The study finds evidence that knowledge diversity has opposing effects on the two creativity components and that user-entrepreneurs (compared with professionals) are better at leveraging diversity for novel game designs.

[Lutter and Weidner, 2021]

In their study "Newcomers, betweenness centrality, and creative success: A study of teams in the board game industry from 1951 to 2017," Lutter and Weidner employ network analysis to understand how network characteristics such as betweenness centrality and the number of newcomers affect the success of teams. This research is based on brokerage and core-periphery structure theories within knowledge diffusion. Creative success is operationalized as usefulness, measured through ratings and the number of awards or nominations received. The study uses regression to analyze the covariates of team networks (a two-mode network of designers and artists connected to games) on creative success, controlling for team characteristics (size, experience) and game characteristics (difficulty, complexity, reimplementations). Findings indicate curvilinear effects of betweenness centrality on success and negative effects associated with the number of newcomers per team.

[Cortinovis and van der Wouden, 2021]

Cortinovis and van der Wouden investigate the impact of prior collaborations on an individual's creative performance as one of the consequences of team composition. The research focuses on factors such as the quality of collaborators, cognitive diversity, status, and experience.

The study is grounded in theories of knowledge diffusion and recombinational creativity. To operationalize creativity, the researchers measure creative performance improvement through game ratings, both categorical and continuous. The method employed in this study involves a 1-1-2-1 matched-samples (of solo projects followed by collaborations followed by a solo project) design for pre- and post-collaboration comparisons, controlling for variables such as the status and experience of the designers. The findings reveal that collaboration with higher-quality designers improved individual performance, possibly due to learning effects. Additionally, there is a nonlinear (U-shaped) relationship with diversity: both low and high overlap in expertise among team members increased the probability of performance improvement. This study highlights the importance of collaborator quality and the complex role of cognitive diversity in creative performance within the board game industry.

These studies explore creative collaborations in the board game industry, emphasizing the roles of team composition in fostering innovation and creative success. Each study focuses on different aspects:

- **[Wachs and Vedres, 2021]** – examines the impact of crowdfunding on innovation within the board game industry.
- **[Pollok et al., 2021]** – investigates how different types of teams, such as user teams versus professional teams, translate diversity into creativity.
- **[Lutter and Weidner, 2021]** – analyzes the effects of network characteristics like betweenness centrality and the presence of newcomers on team success.
- **[Cortinovis and van der Wouden, 2021]** – explores how prior collaborations influence individual creative performance.

These studies employ varying concepts of creativity and innovation. However, the supposed explanatory mechanisms in these studies are not directly observed, e.g., motivational processes, learning, etc. They often control for temporal and relational confounders instead of leveraging this information through a more data-driven approach. For example, studies 1-3 largely ignore team tenure and previous collaborations, while study 4 applies very restrictive sampling.

Moreover, these studies usually do not address the intricate dynamics of team formation, including changes in team composition and self-selection processes. We aim to address these gaps by applying a recent data-driven method, the Relational Hyper Event Outcome Model [Lerner and Hâncean, 2022], to explore the dynamics of team formation and its effects on creative outcomes.

2.7 Research Goals

The primary aim of this work is to explore network effects in team design and their influence on creative outcomes within the board game industry. Specifically, we focus on the dynamics of team formation and the factors that contribute to creative success.

2.7.1 Research Questions

In pursuit of our research goals, we investigate the following questions:

- **RQ1:** What are the main drivers of team formation?
- **RQ2:** How do factors that drive team formation/selection relate to team creative success?

2.7.2 Research Hypotheses

To address these research questions, we formulate and test the following hypotheses within the relational event modeling framework:

- **H1:** Specific hyperedge statistics—a measure used to characterize teams by their past social interactions—increase the likelihood of collaboration. We will explore which hyperedge statistics, as motivated by the literature, contribute to enhanced collaboration.
- **H2:** The probability of collaborating within a team is positively related to the probability of achieving better creative outcomes. We will investigate whether the factors that increase the likelihood of collaboration also lead to higher creative success.
- **H3:** High-creativity teams form differently from low-creativity teams. We will compare the group formation processes, analyzing adverse selection and identifying the characteristics that contribute to the success of high-creativity teams.

By systematically exploring these questions and testing these hypotheses, we aim to provide a comprehensive understanding of the mechanisms driving team formation and their impact on creativity in the board game industry.

Chapter 3

Methods

3.1 BoardGameGeek Dataset

BoardGameGeek (BGG), a community-run database and forum, provides a rich source of observational data on creative collaborations in the board game industry. This data offers a structured way to compare board games and has been extracted repeatedly by researchers (see 2.6). [Shepherd, 2024], who regularly scrapes and compiles the data to power his board game recommender system [1], provided us with unlimited access upon request.

Category	Count	Details
Games	119,392	User-maintained records of game publications
Users	508,901	BoardGameGeek user information
Ratings	59,624,141	Ratings with date, user, and rating information

Table 3.1: Summary of BGG Dataset

The study of possible interactions with unfolding user evaluations is highly interesting but goes beyond the scope of this work. Since we are interested in team formation and performance, we will focus on the detailed information at the board game level, which includes game identification, publication information, game type, contributors, audience/player-related information, game characteristics, popularity and engagement, complexity and accessibility, and game implementations (see Appendix B.1 for full details).

[1] Board game recommender system by markus sheppard `https://recommend.games/`

Visual Exploration of Game Collaborations

BGG (BoardGameGeek) was founded in 2000 by Scott Alden and Derk Solo contributors [2024]. The logging of past and future games into its database is a continuous community project, and today, the platform strongly resembles the structure of a wiki.

We will now visually explore game collaborations, considering a game as a collaboration event. Next to other attributes, a game is associated with a year of publication and a set of designers, encoding the participants and the date of collaboration as well as a rating on the product of their collaboration. It also includes a set of mechanics used by a specific game.

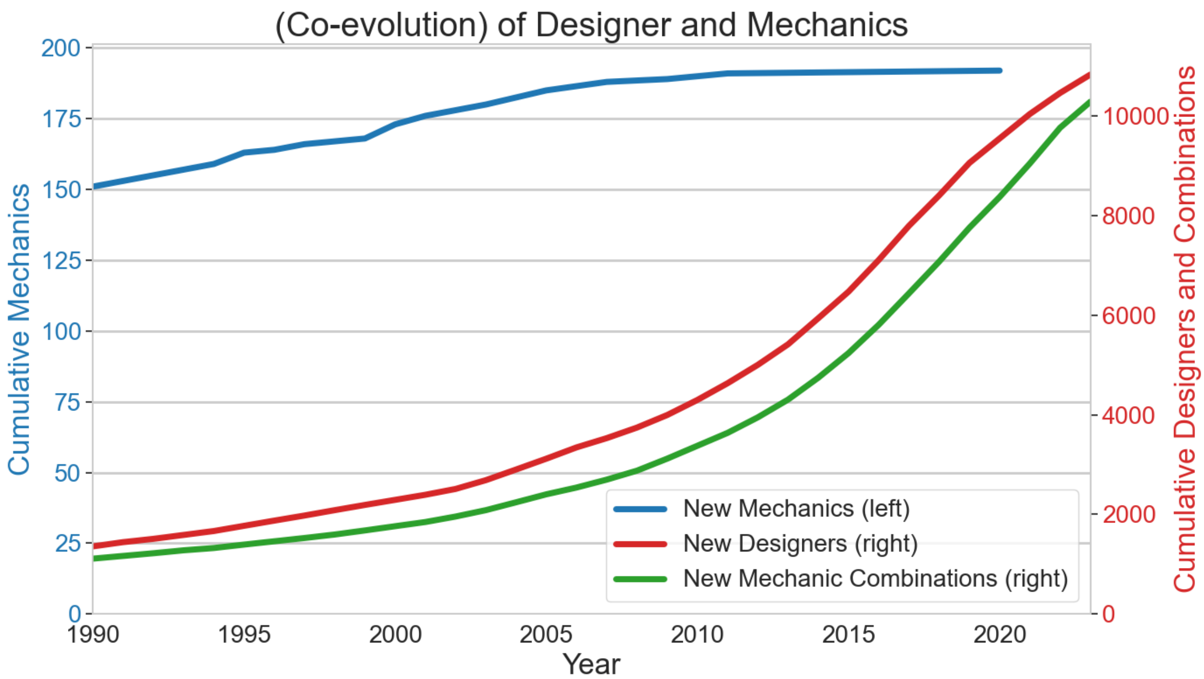

Figure 3.2: Evolution of the Number of Designers, Mechanics, and Mechanics Combinations

Figure 3.2 illustrates the trends in game mechanics combinations over the years with the growing community of BGG designers, displaying both the total number of mechanics and cumulative combinations annually from 1990 to 2023. The data reveals insights into the evolution of game mechanics alongside the size of the game designer community.

From the 1950s to the late 1980s, the number of new mechanics combinations grew slowly (not displayed in the graph but can be inferred by low numbers at the start), indicating a relatively stable period with gradual innovation. Around 1990, the rate of new mechanics combinations began to increase more rapidly. This coincides with the spread of the game design stream Eurogames (see earlier section) in which games are built around mechanics that facilitate social play. This period marks the beginning of a significant growth phase in game mechanics innovation. The total number of mechanics combinations shows an exponential increase, especially from 2010 onward. This superlinear growth continues, reflecting the burgeoning complexity and diversity in game design during these decades. The rapid rise in the total number of mechanics combinations indicates a trend towards more intricate and varied game mechanics being developed and incorporated into board games.

This growth underscores the dynamic nature of the board game industry, with designers continuously exploring and implementing new mechanics to enhance gameplay experiences. The growth in new combinations underscores the innovative spirit and creativity driving the board game industry, with an ever-expanding repertoire of game mechanics contributing to the diversity of modern board games. Alongside this, the total number of designers has steadily increased superlinearly to the present day. How this relates to teams is illustrated in the next graph. The graphic suggests that the co-evolution of the community and games is interrelated. Key elements of our research at first glance seem connected, which we will continue to explore in this work.

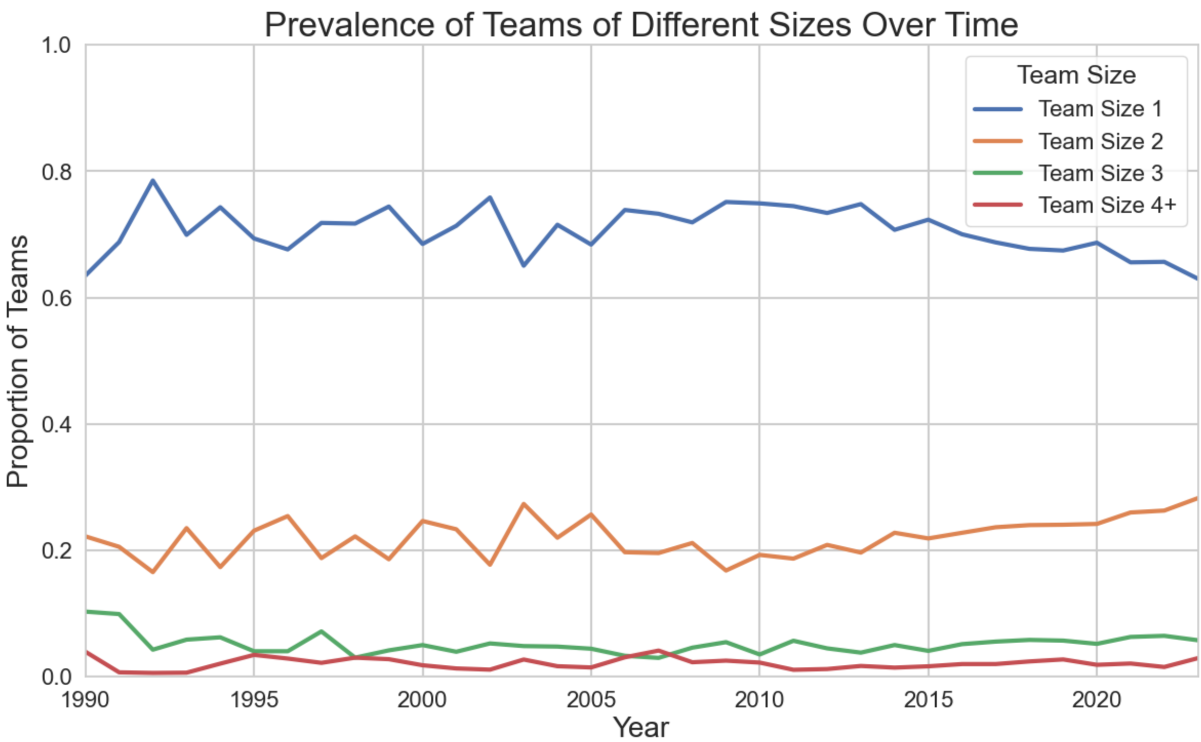

Figure 3.3: Prevalence of Different Team Sizes Over Time

Figure 3.3 shows the prevalence of different team sizes over time from 1990 to 2023. The data reveals several key trends:

1. **Team Size 1 (Solo Designers)**: The proportion of solo designers remains the highest among all team sizes but has shown a slight decline over the years. Despite this, solo design remains a dominant mode of game creation.

2. **Team Size 2**: Teams of two designers have gradually increased in prevalence. This suggests a growing trend towards collaborative efforts between pairs of designers.

3. **Team Size 3/4+**: The proportion of larger teams of three or more is relatively low and stable over the period, indicating that this team size is less common compared to solo or duo teams.

In summary, there is a wealth of quality data on game collaborations, encompassing both the timeline and the designers involved, in a field that has seen significant growth. Games are accompanied by structured data, including detailed mechanics, and are also subject to community evaluations (ratings).

3.2 Data-Driven Analysis of Teams

Building on these insights, the next section will draw from methods in the science of team science[2], particularly focusing on coauthor network analysis, to further explore collaborations of board game designers.

> *" [Models. . .] are not restricted to this application area but can be applied to other situations in which relational hyperevents represent a team that tackles a given task, provides a service, or produces a product and where these events are associated with a measurable outcome. "*
>
> *—Lerner and Hâncean [2022]*

The *science of team science* Liu et al. [2023] studies the dynamics of scientific collaboration. Scientific co-author analysis provides a plethora of metrics that capture the output (e.g., number of published papers) and impact (e.g., number of citations gathered by published papers) of scientific teams. According to a recent review, three factors have accelerated progress in the science of team science:

[2] A new interdisciplinary field that empirically examines the processes by which large and small scientific teams, research centers, and institutes organize, communicate, and conduct research

1. **Data**: Research databases now house millions of articles, grant proposals, patents, and more, providing a detailed view of scientific activity. This extensive repository of data allows researchers to track and analyze the evolution of scientific research over time.

2. **Metrics**: Leveraging this vast amount of data, researchers have created new metrics to quantify scientific activities. These innovative metrics enable the evaluation of theories that were previously considered difficult to measure (e.g., testing and finally rejecting "the common belief that young scientists typically drive breakthroughs in science," as the data suggests success at middle age Liu et al. [2023]). By developing these measures, scholars can better understand the dynamics of scientific productivity and impact.

3. **Method**: Advances in data science, network science, artificial intelligence, and econometrics have equipped researchers with powerful tools to study relationships, predict outcomes, and assess science policy. These analytical methods allow for a deeper examination of the intricate connections within scientific research, making it possible to forecast trends and evaluate the effectiveness of various scientific policies.

Together, these advancements in data collection, metric development, and analytical methods have provided new insights into the dynamics of scientific progress. They enable rigorous testing of established theories, identification of factors that drive productivity, prediction of research outcomes, and formulation of policies to foster scientific innovation.

3.2.1 From Scientific Research to Board Game Innovation

In our effort to study board game designer collaborations, we argue that while the context might differ, similarities in the data-generating processes allow us to repurpose metrics and methods. Specifically, team science methods (e.g., studying patterns in the formation of scientists) can be readily applied to teams in other contexts, provided the data traces have similar characteristics. In Table 3.4, we highlight these similarities.

3.2.2 BGG Data Traces–Analytical Potential

RQ1 concerns the factors explaining team formation, specifically how designers self-select into teams. The factors driving collaboration are multifaceted and complex. Previous work on BoardGameGeek (BGG) has either overlooked the analytical potential of network effects or attempted to control for them rather than utilizing them for deeper analysis.

- Lutter and Weidner [2021] explores the effects of network centrality on creative success by using a two-mode network (game-designer). However, this study fails to apply realistic

Aspect	Description
Collaboration	Participation is optional and voluntary.
Completed Work	Work can be completed alone or in a team, culminating in a publication.
Career Publications	Typically involves producing multiple publications over the course of a career.
Publication Teams	Teams change over time, with designers working solo or in varying teams.
Outcome Similarity	Both collaborations in research and games produce measurable outcomes which represent community evaluations (e.g., ratings, citations).

Table 3.4: Similarity of Collaborations in Scientific Research and Board Game Innovation

null models. For a hypothesis on how the network centrality of a team affects success, it is crucial to control for alternative explanatory variables such as mere activity level or previous success of team members.

- Cortinovis and van der Wouden [2021] investigates how collaborating with more experienced designers impacts an individual's subsequent performance. The study acknowledges the challenge of disentangling temporal and relational effects. To address this, it employs a strict matching process, which results in a small subset of the overall dataset being analyzed, thus limiting the study's external validity.
- Pollok et al. [2021] examine how user innovators are better at turning knowledge diversity into novelty. However, they do not explicitly address knowledge transfers and combinations. For instance, they do not observe whether the knowledge diversity of designers is effectively combined in the games they jointly design. While their reasoning does not necessitate this observation, it could be investigated.

Although not all studies directly focus on network effects, many include variables related to hyperedge statistics and other network characteristics, which can be difficult to interpret in isolation. These limitations underscore the need for more sophisticated analytical approaches that fully leverage network data to study collaborations and their outcomes.

In addition to the static network effects discussed in 2, which influence team formation and performance, dynamic factors also play a crucial role. These factors include prior success, expertise diversity, and temporal dynamics, all of which contribute to the complex nature of collaborative networks.

Prior success refers to a designer's historical performance in terms of game ratings or awards. Designers with a track record of successful games are likely more attractive collaborators, influencing team formation. The hypothesis is that successful designers will continue to attract new partners, leveraging their past achievements.

Experience and expertise diversity within a team refers to the variety of skills, knowledge, and expertise present. For example, combining designers who specialize in different areas, such as mechanics design, thematic storytelling, or graphic design, can lead to more innovative and successful games. The hypothesis here is that teams with diverse expertise are better equipped to produce high-quality and innovative board games.

Temporal dynamics address the influence of time on network formation and evolution. This includes long-term partnerships versus new collaborations and the frequency of collaborations over time. The hypothesis is that the timing of collaborations and the evolution of relationships impact the success and stability of teams, with established partnerships potentially being more stable and productive.

These are only a subset of possible effects that work in tandem and often correlate. In the following section, we will introduce a model framework that allows us to relate the factors that increase co-publication rates to scientific impact, which we will leverage to address RQ2.

3.2.3 Why Do We Need Statistical Network Models?

To empirically estimate network effects on team formation, it is essential to identify patterns in the data that increase the probability of collaboration. BoardGameGeek (BGG) data can be understood as relational event data, incorporating both temporal and relational dependencies that influence the likelihood of hypothetical teams forming. The modeling task is to estimate the interaction rate of all possible teams based on past data.

Examples of Factors Increasing Collaboration Likelihood:

- **Previous Collaboration**: Teams are more likely to form if the members have worked together before.
- **Recent Collaboration**: Recent collaborations between members increase the probability of future teamwork.
- **Popularity of Successful Members**: A successful team member may be more popular and sought after for collaborations.
- **Member Activity**: Higher activity levels of members increase their chances of being selected for a team.
- **Similar Past Projects**: Members who have worked on similar projects in the past are more likely to collaborate.

- **Shared Connections**: Members who share a mutual developer friend, even if they have never worked together, are more likely to form a team.

Since this kind of data is only observational, causality cannot usually be established. Nevertheless, when trying to establish a significant correlation with this type of data, it is important to rule out alternative explanations. For instance, if a homophily hypothesis suggests that two developers with similar prior success are more likely to become part of a team, one needs to also check for alternative explanations such as how often they publish at all. Maybe they also collaborate because they have worked on similar projects in the past, and so on. An appropriate statistical model allows us to account for multiple effects simultaneously, rule out alternative explanations, and capture dependencies between explanatory variables sufficiently.

More formally, the data has the following characteristics that need to be taken into account:

Data Characteristics

- **Temporal**: Each event occurs at a specific point in time, providing a temporal context to the activities of the actors involved.
- **Relational**: Events are defined by the participation of actors, highlighting the connections and interactions between them.
- **Polyadic**: Events can involve more than two actors, with a minimum of one actor, allowing for the analysis of complex interactions.
- **Outcome**: Each event can be associated with measurable outcomes, such as ratings or citations, enabling the evaluation of its impact.

Nature of Polyadic Interaction Data:

- Go beyond social network data from technology-mediated interactions.
- Time-stamped polyadic interaction processes are common across various empirical research areas.

Examples of Undirected Polyadic Interaction Networks:

- Co-authors publishing a paper together.
- Coordination in task-oriented teams.

- Writers collaboratively working on a script.
- Developers working on a software project as a team.
- Meetings with multiple participants.

Examples of Directed Polyadic Interaction Networks for Board Game Designers in a Team Creativity and Organizational Research Context:

- A lead designer explaining the game concept to the design team.
- A creative director presenting a strategy or vision for the game's development to the team.

3.3 Model Framework

Given polyadic interaction data with the previously discussed characteristics, our research investigates group formation and how these formations relate to respective outcome variables. At the core is the question of identifying generalizable patterns that increase the probability of future collaboration based on a history of events. Relational Hyperevent Models (RHEM) analyze social interactions by using covariates for groups of participants, addressing the limitations of traditional models that focus on pairs of individuals. We also introduce the related Relational Hyperevent Outcome Model (RHOM), which allows for modeling team reflexivity with respect to collectively achieved outcomes.

Disclaimer: The measures and definitions presented in this section are adapted from existing methodologies as described in the following works:

- Hâncean et al. [2020] - The coauthorship networks of the most productive european researchers
- Lerner et al. [2021] - Dynamic network analysis of contact diaries
- Lerner and Hâncean [2022] - Micro-level network dynamics of scientific collaboration and impact: Relational hyperevent models for the analysis of coauthor network
- Lerner and Lomi [2023] - Relational hyperevent models for polyadic interaction networks
- Bright et al. [2023] - Investigating the dynamics of outlaw motorcycle gang co-offending networks: The utility of relational hyper event models

3.3.1 Background

Relational Event Models (REM) specify collaboration probabilities for dyadic interactions. RHEM extends the dyadic model to capture interactions within groups of various shapes and structures, overcoming the limitations of traditional sender-receiver models. Before defining RHEM, it is essential to introduce the underlying mathematical concepts.

Hypergraphs

In graph theory, a graph is defined as $G = (V, E)$, a pair consisting of a set of vertices (nodes) and a set of edges (links) [3]. An edge connects two nodes. Edges can be:

- **Directed or undirected**: Representing ordered pairs (tuples) or unordered pairs (sets), respectively. For example, directionality distinguishes between meeting events or sending messages from A to B.
- **Weighted or unweighted**: Indicating different strengths of relations or interactions.
- **Multi-modal or single-mode**: Capturing interactions between different types of nodes, such as games and game developers, or interactions within a single type of node.

Moving from graphs to hypergraphs is achieved by replacing edges with hyperedges. A hyperedge is a generalization of an edge, capable of modeling polyadic interactions. Simply put, a hyperedge can connect more than two nodes. The following defines the unweighted, single-mode generalizations for both undirected and directed graphs:

- **Undirected hypergraph**: A pair $G = (V, H)$, where V is a finite set of nodes and $H \subseteq P(V)$ is a set of subsets of V. An element $h \in H$ is called a hyperedge.
- **Directed hypergraph**: A pair $G = (V, H)$ where V is a finite set of nodes and $H \subseteq P(V) \times P(V)$ is a set of pairs of subsets of V (hyperedges). A hyperedge $h = (u, v) \in H$ consists of a source set $u \subseteq V$ and a target set $v \subseteq V$.

[3] The terms vertices and edges or nodes and links are used interchangeably. The referenced work typically uses mixed terminology, so for simplicity, we adopt the mixed terminology of (hyper)edges and nodes. Throughout this work, we will interchangeably use vertices/nodes/actors and edges/links/relations.

Hyperevents

In our research, we view a network as a series of timestamped interaction events instead of static relational states (for theoretical motivation, refer to 2).

A hyperevent is an extension of a hyperedge, including a timestamp and an event type. Specifically, a hyperevent is defined as:

$$e = (t_e, h_e, x_e)$$

where:

- t_e is the timestamp of the event.
- h_e is the hyperedge capturing the relational aspect.
- x_e is the event type.

If we assume only a single event type, the hyperevent simplifies to:

$$e = (t_e, h_e)$$

A network can then be represented as a sequence of these past events, or hyperevents:

$$E = (e_1, \ldots, e_n)$$

3.3.2 Relational Hyper Event Model (RHEM)

Intuition

RHEM: Which hyperedge statistics increase the probability of an edge occurring?

In essence, the intensity, or event rate, $\lambda(t_e, h)$, represents the expected number of papers that h will co-publish within a unit time interval beginning at t_e.

Definition

Consider a sequence of polyadic events with timestamps, represented as

$$E = (e_1, \ldots, e_N),$$

where each event e_i is described by (h_i, t_i, x_i).

Define the set of all events occurring before time t (past events) as

$$E_{<t} = \{e_i \in E \mid t_i < t\}.$$

We assume that events are conditionally independent given the prior sequence of events.

Relational (hyper)event models define the probability distribution for the sequence as follows:

$$P(E) = \prod_{i=1}^{N} P(e_i \mid E_{<t_i}).$$

The conditional probabilities $P(e_i \mid E_{<t_i})$ are modeled using the hazard or intensity functions $\lambda(t, h)$. Here, $\lambda(t, h)$ represents the expected number of events per time unit for the hyperedge h. For a point in time t, let $E[< t]$ denote events e in E occurring before t. For each subset h of actors A, the hazard rate $\lambda(t, h)$ is defined as:

$$\lambda(t, h) = \lim_{\Delta t \to 0} \frac{N_{t+\Delta t}(h) - N_t(h)}{\Delta t}$$

where $N_t(h)$ is the number of events involving subset h up to time t.

Typically, the hazard function is specified as a parametric function of statistics $s = (s_1, \ldots, s_k)$:

$$\lambda(t, h) = \exp\left(\sum_{j=1}^{k} \theta_j \cdot s_j(h; E_{<t})\right).$$

Hazard Specification

The hazard rate, which approximates the expected number of events per time unit on h at time t, decomposes into a global baseline hazard $\lambda_0(t)$ and a relative hazard $\lambda_1(t, h)$:

$$\lambda(t, h) = \lambda_0(t) \cdot \lambda_1(t, h)$$

The relative hazard $\lambda_1(t, h)$ is specified in parametric form as:

$$\lambda_1(t, h) = \exp\left(\sum_{j=1}^{k} \theta_j \cdot s_j(t, h)\right)$$

Here, $s_j(t, h)$ are explanatory variables (hyperedge statistics) quantifying characteristics that explain the collaboration rate. These statistics can be actor-level characteristics or functions of the event history $E[< t]$, explained in detail in Section 3.3.4. The parameters θ_j indicate the impact of these statistics on the event rate. If θ_j is positive, the rate for h is higher; if negative, the rate is lower; and if zero, the statistic has no effect.

The hyperedge statistics can operationalize hypothetical effects explaining multi-actor interactions. The associated parameters, estimated by the maximum likelihood of the partial likelihood function[4], can be used for statistical testing of respective effects.

Modeling Relative Event Rates with CoxPH

The model variant used here only models the relative event rate and not the full event rate, specifying relative event rates for each hyperedge up to a global base rate but not exactly when or how often this is the case. CoxPH uses partial likelihoods, a statistical method to estimate the hazard parameters without specifying the full model.

The partial likelihood of a hyperedge h_i is given by:

$$P(h_i \mid E_{<t_i}; \theta) = \frac{\lambda_1(t_i, h_i, \theta, E)}{\sum_{h \in R_{t_i}} \lambda_1(t_i, h, \theta, E)}$$

where the hazard function $\lambda_1(t_e, h_e, \theta, E)$ represents the intensity of the event e given the parameters θ and the sequence of past events E. The partial likelihood compares the relative event rate of the hyperedge that experienced the event to the sum of all hyperedges' relative event rates that could experience an event at time point i. The set of all hyperedges that could receive an event is called the Risk Set.

The likelihood function $L(\theta)$ for the full sequence of events can be given without knowledge of the base event rate:

4 The global baseline hazard $\lambda_0(t)$ could be estimated by non-parametric functions, but in our case, an approach is used which does not require full model specification.

$$L(\theta) = \prod_{e \in E} \frac{\lambda_1(t_e, h_e, \theta, E)}{\sum_{h \in R_{t_e}} \lambda_1(t_e, h, \theta, E)},$$

where the sum in the denominator iterates over all hyperedges h in the risk set R_{t_e}. The Risk Set is the set of all hyperedges that could have experienced an event at the respective time point but did not.

This way, parameters θ can be estimated by maximizing the partial likelihood over all observed events:

$$\hat{\theta} = \arg\max_{\theta} \prod_{i} P(e_i \mid E_{<t_i}; \theta)$$

Risk-Set Specification

The number of hyperedges that could receive an event is usually astronomically large, which is why relational hyperevent models use different variants for specifying the risk set. Here are the main approaches:

- **Full Risk Set**

  The full risk set includes all possible hyperedges:

  $$R_{t_e} = \{h \subseteq V\}$$

  Assuming a network of 1000 actors and a maximum edge size of 10, the number of possible subsets can be calculated as:

  $$\sum_{k=1}^{10} \binom{1000}{k} \approx 2.702882 \times 10^{20}$$

  This number is computationally intractable due to its large size.

- **Conditional-Size Risk Set**

  In practice, the conditional-size risk set is typically used:

  $$R_{t_e} = \{h \subseteq V \mid |h| = |h_e|\}$$

  However, this approach does not significantly reduce the order of magnitude when considering larger hyperedges. Nevertheless, it is usually applied because hyperedge size often strongly correlates with event intensity. Restricting the risk set to hyperedges of equal size helps to avoid this potentially confounding effect. [5]

- **Case-Control Sampling**

  Case-control sampling is generally applied to sample non-events from the risk set, making the computation more manageable. For each event, the sampled risk set $\widetilde{R}_{t_i}$ is constructed as follows:

  - It includes the event that occurred at time t_i (referred to as the case).
  - Additionally, m non-events (controls) are selected at random, uniformly and independently, from the full risk set R_{t_i}.

  Using this sampled data for estimating the CoxPH model yields a consistent estimator.

3.3.3 Relational Hyper Event Outcome Model (RHOM)

Intuition

Relational outcomes serve two purposes in modeling: they act as explanatory variables for future events and as response variables for joint event outcomes. This allows for the extension of RHEM with new hyperedge statistics based on past outcomes and enables RHOM to predict event outcomes.

RHEM predicts the probability of events for all hyperedges in a risk set, specifying event rates or publication intensities. In contrast, RHOM focuses on hyperedges that experience events, predicting outcomes such as average ratings.

In RHOM, explanatory variables for the relational outcome y_e can be functions of previous events. For example, the success of a board game (measured by its average rating) can be influenced by the developers' past collaborations and their previous games' performance.

Definition

Both RHEM and RHOM model sequences of board game development events $E = (e_1, \ldots, e_N)$, where each event $e \in E$ is a tuple:

$$e = (t_e, h_e, x_e, y_e)$$

[5] Lerner describes the conditional size model as: "(groups of) actors competing for participation in events"

Key Differences:

RHEM specifies the relative rate $\lambda_1(t, h)$ for all hyperedges h in the risk set.

RHOM specifies conditional probability distributions $f(y_e \mid t_e, h_e)$ for the relational outcome y_e.

Relational outcomes can be any measurable result, such as the average rating, number of votes, or rank, and can be quantified as a binary (success vs. failure), ordinal, or numeric response variable. For numeric response variables, it is useful to normalize the outcome variable:

$$y_e = c_e - \frac{\sum_{e \in E : x_e = x_e \wedge t_e = t_e} c_e}{|\{e \in E : x_e = x_e \wedge t_e = t_e\}|}$$

Here:

- c_e denotes the observed average rating of the board game represented by e.
- x_e represents the genre of the board game.
- t_e denotes the year of publication.

The normalized outcome y_e is positive (negative) if the board game e receives a better (worse) rating than the average game from the same genre published in the same year.

This metric can then be used inside hyperedge statistics to condition the event rate on prior success (see the next section) or as a response variable.

RHOM specifies the likelihood of an observed sequence of relational hyperevents E by:

$$L(\theta) = \prod_{e \in E} f(y_e \mid t_e, h_e, \theta, G[E, t_e])$$

where f represents a distribution for the relational outcome y_e, typically chosen from the family of generalized linear models (GLM).

3.3.4 Hyperedge Statistics for RHEM and RHOM

Hyperedge statistics $s_j(t, h)$, as discussed in Section 3, are essential for both RHEM and RHOM.

These statistics:

- Describe hyperedge characteristics.
- Serve as covariates to predict relative differences in event rates.
- Use MLE of θ_j to determine their impact on event rates:
  - $\theta_j > 0$: Increases rate.
  - $\theta_j < 0$: Decreases rate.
  - $\theta_j = 0$: No effect.
- Reflect network effects, as hyperedges are part of a past event network.
- Enable statistical testing of hypotheses regarding these network effects.

Covariates are grouped into:

- Actor-level characteristics (attribute effects).
- Functions of event history $E[< t]$ (network effects).

The underlying network defines the scope of possible statistics. While an exhaustive list for all network types (undirected, directed, single-mode, two-mode, weighted, unweighted) is beyond this scope, our focus is on undirected hyperedge events, which are suitable for meeting events. We will now define the basic structures for these events.

Decay in the Influence of Past Events Over Time

Given a sequence of publication events

$$E = \{(t_1, j_1, I_1, J_1), \ldots, (t_n, j_n, I_n, J_n)\}$$

, the value of network attributes at time t is a function of earlier events:

$$E_{<t} = \{(t_m, j_m, I_m, J_m) \in E \mid t_m < t\},$$

where the effect of past events decays over time with a given half-life period $T_{1/2} > 0$. This decay is modeled using the following weight factor, which depends on the elapsed time $t - t_m$:

$$w(t - t_m) = \exp\left(-\frac{[t - t_m]\log 2}{T_{1/2}}\right).$$

While using a decay is not strictly necessary, it is theoretically plausible that the effect of past collaborations or citations diminishes over time. Incorporating a decay factor often leads to a better model fit.

Network Effects

Network effects characterize hyperedges based on the history of the process, that is, on sequences of previously observed events. We will now define the main network effects for undirected (unweighted, one-mode) hyperedge events:

Exact Repetition

Repetition or familiarity effects are simple hyperedge statistics that count the repetitions of a hyperedge in the event history.

The repetition statistic associated with a given hyperedge h at time t counts the number of previous events in $E_{<t} = \{e = (t_e, h_e) \in E; t_e < t\}$ whose set of participants h_e is identical to h, weighted by the respective decay factor. Formally, it is defined by

$$\text{repetition}(t; h; G[E; t]) = \sum_{e \in E_{<t}} w(t - t_e) \cdot \mathbb{I}(h = h_e),$$

where $\mathbb{I}$ is the indicator function that is one if the argument is true and zero otherwise.

Subset Repetition

Exact repetition is very limited and cannot identify dense subgroups within a whole. Subset repetition can capture aspects of participant selection based on familiarity of varying orders.

Formally, subset repetition is defined in two steps. First, the hyperedge degree $\deg(t, h; G[E; t])$ measures how many board games have been co-developed by all members of h (potentially in collaboration with other game developers outside of h) before time t:

$$\text{hy.deg}(t, h; G[E; t]) = \sum_{e \in E_{<t}} \mathbb{I}(h \subseteq h_e),$$

(The indicator function $\mathbb{I}$ is one if the argument is true and zero otherwise.)

For a given integer $p \in \mathbb{N}$ (specifying the size of subsets to be repeated), subset repetition of order p is defined by:

$$\text{sub.rep}^{(p)}(t, h, G[E;t]) = \sum_{h_p \in \binom{h}{p}} \text{hy.deg}(t, h_p, G[E;t]) \cdot \frac{1}{\binom{|h|}{p}},$$

where $\binom{h}{p} = \{h_p \subseteq h : |h_p| = p\}$ denotes all subsets of h of size p. In other words, subset repetition of order p is the average hyperedge degree over all subsets of size p of the focal hyperedge h.

For the special case that $p = 1$, subset repetition is equivalent to repetition and is a measure of individual publication activity. Positive parameters with respect to this statistic can indicate a tendency for preferential attachment.

Closure

The closure statistic measures the extent to which members of a hyperedge h have co-participated in previous events with common third actors w. Unlike subset repetition of order three, these third actors w can be outside the focal hyperedge h, and different members of h may have co-participated with w in different past events. Formally, closure is defined by

$$\text{closure}(t; h; G[E;t]) = \frac{1}{\binom{|h|}{2}} \sum_{\{u,v\} \in \binom{h}{2} \wedge w \neq u,v} \min[\text{hy.deg}(t; \{u, w\}), \text{hy.deg}(t; \{v, w\})]$$

In the formula above, we iterate over all combinations $(\{u, v\}, w)$ such that $\{u, v\} \in \binom{h}{2}$ is an unordered pair of actors who are both members of the focal hyperedge h and w is any actor different from u and v. For each of these triples, we compare the hypergraph degrees of the two sets $\{u, w\}$ and $\{v, w\}$ and add up the minimum of these two values. The resulting sum is divided by the number of unordered pairs within h.

Prior Outcome as Covariate

RHOM models outcomes and includes the effect prior outcomes have on event rates. Assessing outcome effects can be achieved by defining subtypes of subset repetition. The prior success/outcome measures are adapted from existing work and defined in a two-step process. First, cumulative prior joint performance is gathered, which is then used to specify either prior joint successes or a team's disparity with respect to prior successes.

To assess the **cumulative prior joint performance** of groups of game developers $h \subseteq V$ of any size, we sum the relational outcome y_e over past board game developments co-authored by all members of h (potentially including other developers outside of h):

$$\text{performance}(t, h; G[E; t]) = \sum_{e \in E_{<t}} y_e \cdot \mathbb{I}(h \subseteq h_e).$$

Based on this measure of prior joint performance, we define the **prior success of order** p of a hyperedge h by iterating over all subsets $h_p \subseteq h$ of size p and summing the cumulative prior joint performance of these subsets h_p. This measure is normalized by the cumulative degree of all those subsets, leading to the statistic[6]:

$$\text{prior.success}^{(p)}(t, h, G[E; t]) = \frac{\sum_{h_p \in \binom{h}{p}} \text{performance}(t, h_p, G[E; t])}{\sum_{h_p \in \binom{h}{p}} \text{hy.deg}(t, h_p, G[E; t])}.$$

To assess how much members of a development team differ in terms of their prior success, we calculate the standard deviation of the individual performance of team members. In formulas, we abbreviate performance$(t, \{v\}, G[E; t])$ by performance(v) and write $\bar{p}(h)$ for the mean performance:

$$\bar{p}(h) = \frac{\sum_{v \in h} \text{performance}(v)}{|h|}$$

then,

$$\text{success.disparity}(t, h, G[E; t]) = \sqrt{\frac{1}{|h| - 1} \sum_{v \in h} [\text{performance}(v) - \bar{p}(h)]^2},$$

where we set the success disparity of hyperedges of size one to zero.

[6] Practically, values less than 0 can be set to 0, which simply counts the successes and not the failures accordingly, as suggested by the authors

Attribute Effects

Beyond their past collaborative activities, hyperedges/teams can be described by summary statistics of attributes at the node/actor level.

Node attributes may be exogenous (such as "age") or endogenous (such as the number of past successes). Since hyperedge statistics of $|h| = 1$ can describe equivalent aspects as endogenous node attributes, we will primarily focus on exogenous node attributes. These attributes can be of binary, categorical, or numerical value. Examples include gender, age, tenure, nationality, roles, or memberships.

Attributes can be formalized as node-level functions, e.g., $x : V \longrightarrow \{0, 1\}$, where x is a time-constant node-level function that assigns a value to each possible node. The definition of time-varying functions is straightforward.

Node attribute statistics provide a summary function of a specified node attribute over the participating nodes. There are a variety of summary functions representing central tendencies or dispersion/heterogeneity over the event participants.

First Order Effects

The aggregation functions `mean`, `sum`, `max`, `min`, and `product` summarize central tendencies of values across the nodes of hyperedges. These aggregated statistics of node attributes can be used to investigate "first order effects." Specifically, they help determine whether hyperedges, characterized by participants, sources, and/or targets with higher or lower attribute values, tend to have different event rates.

Dispersion

The aggregation functions `standard deviation (std)`, `sample standard deviation (sample_std)`, `absolute difference (abs_diff)`, and `categorical difference (cat_diff)` capture the dispersion or heterogeneity of values across the nodes within hyperedges. These statistics help examine whether nodes tend to interact with others that are similar or different from themselves in terms of attribute values. A positive parameter value indicates heterophily (preference for interacting with dissimilar others), while a negative parameter value indicates homophily (preference for interacting with similar others).

Statistical Possibilities with Other Network Structures

As indicated earlier, the underlying network structure defines the range of possible effects. For example, directed networks significantly increase the combinations and variations of subset repetition. However, this is not relevant for our use case, as collaboration in board game development does not reflect directed interaction.

Another distinction involves understanding games as references to game mechanics. In this way, the development of board games, previously understood as an undirected one-mode network, can be conceptualized as a two-mode network of game developers and game mechanics. Although a two-mode network increases the possibilities for defining network effects, it comes at the cost of increased complexity, making the interpretation of network effects more difficult.

In a later step, we will include a two-mode specification, allowing us to explore how designers (or designer teams) have utilized mechanics. While the formal definition of two-mode statistics is beyond the scope of this discussion, they function equivalently. However, since the coevolution of game mechanics and game developers is not immediately related to our research question, we will first focus on investigating the main effects in the collaboration network.

3.4 Model Specification

Descriptive Approaches

Before specifying the models, it is important to clarify that the empirical approach of this work is descriptive rather than causal due to the observational nature of the data. This approach combines elements of exploratory analysis with more classical methods. Here, we clarify our approach:

Mixed Approach:

- **Empirical Regularities and Generalizable Facts:** We aim to establish observational regularities about board game collaborations that can motivate new theory development.
- **Classic Regression:** As part of our descriptive approach, we engage in formal hypothesis testing to examine relationships between variables.

Based on our research questions, our work situates itself between these two approaches. We aim to uncover broad observational patterns (RQ1: What are the main drivers of team formation?) and test specific relationships (RQ2: How do factors that drive team formation/selection relate to team creative success?), as well as compare different formation processes (RQ3: How do high-creativity teams form differently from low-creativity teams?) within the board game collaboration network.

Theory: Formation and Creative Success

Understanding team formation and creative success in board game collaborations involves several theoretical perspectives that can inform the coding of relevant variables:

Psychology-Based Research

Team formation provides the foundation for key psychological processes that drive collaboration and creativity. As discussed earlier in Section 2, recent research has identified the following key aspects of team formation. We will now interpret these in our statistical framework given the BGG dataset.

- **Team Size:**
  - **Motivation:** Affects the diversity of perspectives and the complexity of coordination. Larger teams offer more ideas but face coordination challenges.
  - **Hyperedge Statistic:**
    - *Hyperedge Size*
- **Team Tenure:**
  - **Motivation:** Influences trust and psychological safety. Longer tenure can enhance coordination but may also lead to groupthink.
  - **Hyperedge Statistics:**
    - *Repetition* (number of times the team has collaborated)
    - *Subset Repetition* (familiarity among subsets of team members)
    - *Average Time Since First Published Game* (temporal aspect)
- **Knowledge, Skills, Abilities, and Other Characteristics:**
  - **Motivation:** These characteristics reflect the expertise and capabilities of team members.
  - **Hyperedge Statistics:**

    * *Prior Individual Success* (proxy for skills and status)
    * *Average Portfolio Size* (#mechanics average over team members)

- **Team Diversity:** Organizational Psychology research considers creative processes as balancing between opposing forces. Diversity is quantified as disparity when it emphasizes imbalances and potential conflicts. It is quantified as heterogeneity when it reflects a beneficial range of perspectives and experiences.
  - **Heterogeneity:** Quantifies overall variability within the team using Standard Deviation (SD).
  - **Disparity:** Quantifies inequalities or imbalances between members using Mean Absolute Difference (MAD).
- **Skills-Related Diversity:**
  - **Motivation:** Diversity in skills can enhance creativity by introducing varied perspectives but may also lead to conflict if not managed properly.
  - **Hyperedge Statistics:**
    * *Prior Success Disparity:* Measures inequality in previous individual achievements among team members, highlighting potential imbalances.
    * *Past Portfolio Size Heterogeneity:* Assesses the variety within team members' previous project experiences, indicating a range of skills and perspectives.
    * *Tenure Disparity:* Examines differences in the length of time team members have been active in the industry, reflecting varying levels of experience and potential power dynamics.
- **Background-Related Diversity:**
  - **Motivation:** Diversity in demographic backgrounds can bring varied perspectives and experiences but may also lead to conflict if not managed properly.
  - **Hyperedge Statistics:**
    * *Gender Disparity:* (inferred from first names): Measures gender balance within the team, highlighting gender-based inequalities and their influence on team dynamics.

Sociology-Based Research

Team formation and creative success are often discussed in terms of homophily, the tendency to associate with similar others. This discussion also implicitly includes heterophily, the tendency to connect with dissimilar individuals, as its counterpart. In addition to homophily, we consider how teams are embedded into the network of past events, which encodes relationship strength, resource/information access, and status effects. Due to the overlapping nature of these constructs, we group the statistics based on two primary aspects: homophily, which captures the diversity and similarity within teams, and embeddedness, which captures how team members' past interactions and successes influence current dynamics. Lastly, we explore how designers' past usage of mechanics influences their collaborative behavior.

Hyperedge Statistics:

- **Homophily: Diversity and Similarity within Teams**
  - *Gender Disparity*: Measures the variety in team member backgrounds. Relevant for understanding gender-based diversity.
  - *Average Gender*: Average gender composition of the team. Provides insight into the gender balance within the team.
  - *Tenure Diversity*: Measures the variation in team members' lengths of service. Reflects a mix of experience levels.
  - *Success Diversity*: Measures the variation in team members' past successes. Highlights the range of previous achievements.
- **Embeddedness I: Influence of Past Interactions and Successes**[7]
  - *Team Size*: Measures the number of team members. Influences communication dynamics and resource allocation.
  - *Individual Repetition*: Measures the frequency of individual team members working together in the past. Indicates established working relationships and trust.
  - *Individual Success*: Measures the past success rates of individual team members. Reflects the individual track records within the team.
  - *Subset Repetition*: Measures trust within relationships based on repeated collaborations. Highlights subgroups with strong working bonds.
  - *Subset Success*: Measures trust within relationships based on successful past collaborations. Indicates effective subgroups with proven success.
  - *Closure*: Measures the number of structural holes bridged, facilitating information flow and innovative solutions. Indicates the team's ability to connect disparate ideas and knowledge.
- **Embeddedness II: Influence of Past Mechanic Usage**
  - *Portfolio Size avg*: Measures the average experience levels with using varied game mechanics. Indicates creative input.
  - *Portfolio Size heterogeneity*: Measures the variety of experience levels with using varied game mechanics. Indicates diversity in creative input.

[7] All the metrics can be extended to the use of mechanics, which will be seen in the results section. These statistics will encode preferences, habits, and experience with respect to mechanic use.

- *Team-Mechanic Closure*: Measures the extent to which team members have used similar mechanics in the past. As a normalized measure of the joint mechanics the team members have independently used, in the context of homophily it can be interpreted as selecting team members of similar past projects.

In summary, while the exact processes that drive team formation and creative success may not be directly observable, the traces left by these processes can be encoded in data and interpreted within these theoretical frameworks. This approach allows us to infer diverse dynamics from the BGG dataset, investigating patterns that contribute to successful collaborations.

3.5 Data Cleaning and Preparation

Data Cleaning Process

We conducted the following cleaning process:

- Restricting the dataset to games released between 1990 and 2023, as this period is relevant to the Eurogames era, which focuses on mechanics.
- Limiting the number of designers per game to a maximum of four, as over 99% of games have four or fewer designers.
- Ensuring that each game included specified mechanics information, as mechanics form the basis for our novelty measure.
- Excluding primitive expansions:
  - We identified and excluded **1986 primitive expansions** from the dataset. These expansions were versions of previously released games that did not change their core mechanisms.
  - For example, "Lost Worlds"[8] (war book game) has **68 expansions** that primarily add new scenarios and components without altering the fundamental mechanics. Including such expansions would distort the novelty analysis by suggesting these games have higher innovation levels due to their multiple editions.
- Excluding primitive reimplementations:
  - We excluded **1413 primitive reimplementations**. These are games that have been reissued or rebranded with unchanged mechanics.

[8] https://boardgamegeek.com/boardgame/1421/lost-worlds

- An example is the game "Fluxx"[9] (card game), which, despite having numerous reimplementations (totaling **29**), retains its core mechanics across different versions. Including such games would misrepresent their novelty, as the mechanical innovation is minimal despite the various reissues.

Procedure for Name-Based Gender Inference

The BGG dataset lacks gender information, so we follow a procedure outlined in Vedres and Vasarhelyi [2023]. This method only accounts for binary male and female classifications, utilizing a baby names database to infer gender. The specific steps are:

- **Reference Name Database**:
  - Utilized the 2016 baby names database from the US Social Security Administration (SSA[10]), which contains names along with their associated sex and frequency counts.
- **Calculating Probabilities**:
  - Organized names by sex and calculated the likelihood of each name being male based on frequency counts.
- **Handling Unisex Names**:
  - Combined counts for names used for both sexes and adjusted missing values for accurate probability estimates.
- **Applying Probability Thresholds**:
  - Assigned gender labels based on probability thresholds (≥ 0.9 for male and ≤ 0.1 for female), labeling intermediate probabilities as unknown to ensure high precision.

Measuring the Novelty of Mechanics Combinations

Observation at the level of mechanics combinations focuses on how different game mechanics integrate and interact within a board game to create unique gameplay experiences. This involves identifying the core mechanics (such as deck-building, worker placement, or resource management) and examining their interactions.

[9] `https://boardgamegeek.com/boardgame/258/fluxx`
[10] `https://www.ssa.gov/oact/babynames/`

Qualitative aspects, such as how these mechanics influence each other during gameplay, affect player decisions and strategies, and impact overall player enjoyment, cannot be directly observed. Instead, these qualitative aspects are assumed to be reflected in the game's rating. Therefore, mechanic observations are used to code novelty aspects.

Novelty measures, as used in prior research (Wachs and Vedres [2021]; Pollok et al. [2021]), are identified by two questions:

- **Micro Analysis: Is the Combination New?**
  - **Focus**: Compare the observed combination with existing games to identify unique or innovative interactions and aspects that distinguish it from previously seen mechanics.
  - **Metric**: Novelty $N(g_i)$, which measures the minimum Hamming distance[11] between the game g_i and any previously published game:

    $$N(g_i) = \min_{g_j \in G_{today}} Hamming(g_i, g_j),$$

    where G_{today} refers to the set of all other published games up to (including) the focal year. To simplify, a binary variable $N_1^0(g_i)$ can be used, where $N_1^0(g_i) = 1$ if $N(g_i)$ is greater than 0, otherwise $N_1^0(g_i) = 0$. This binary variable indicates whether the combination is entirely new compared to all past games.
- **Meso Analysis: How Typical is the Combination?**
  - **Focus**: Assess the prevalence and commonality of the combination within the board game industry.
  - **Metric**: Distinctiveness $D(g_i)$, which measures the average Hamming distance between the game g_i and all previously published games:

    $$D(g_i) = \frac{\sum_{g_j \in G_{today}} Hamming(g_i, g_j)}{|G_{today}|},$$

    where G_{today} refers to the set of all other published games up to (including) the focal year, and the Hamming distance counts the number of differing mechanics.

In this analysis, we consider only the meso measure for the following reasons:

- The micro measure is too sparse, focusing on whether a combination is entirely new, leading to a binary outcome that lacks the nuance needed for more detailed comparisons.

[11] Hamming distance is used instead of Euclidean distance for binary vectors because it directly counts the number of differing elements. This is ideal for comparing categorical data where each bit represents an equally significant game mechanic.

- The micro measure has been primarily used to compare two distinct groups, which is not the primary focus of our analysis.
- We require a more continuous measure, allowing for a more granular understanding of novelty.

Chapter 4

Results

4.1 Overview

Our analysis seeks to understand the formation and success of design teams within the dynamic and varied context of board game design. The primary research goals are to explore the network effects in team design and their influence on creative outcomes.

RQ1: What are the main drivers of team formation?

We characterize collaborations by hyperstatistics, which capture the composition of actors and their past interactions. To estimate the contribution of these covariates on the relative event rate (the probability of a collaboration taking place given its covariates), we use Cox proportional hazards (CoxPH) survival analysis.

RQ2: How do factors that drive team formation relate to team creative success?

By linking the identified drivers to creative outcomes, we aim to understand whether these factors contribute to higher creativity rates in board game design. This involves analyzing whether the elements that facilitate team formation also correlate with improved creative performance.

To measure creative outcomes, we operationalize creativity through two distinct dimensions: novelty and usefulness. Novelty is assessed by measuring the average Hamming distance between a game's mechanics and those of previously published games. Usefulness is reflected in game ratings, which capture qualitative aspects like the influence of mechanics on gameplay, player decisions, and enjoyment.

Hypotheses Testing

To address the research questions, we propose and test the following hypotheses:

H1: Specific hyperedge statistics, which characterize teams by their past social interactions, increase the probability of collaboration. We build multiple CoxPH models and evaluate which model best fits the data.

H2: The probability of collaborating within a team is positively related to achieving better creative outcomes. We will run OLS regression on the covariates used for CoxPH to see how factors that drive collaboration relate to performance.

H3: High-creativity teams form differently from low-creativity teams. We will compare the 25th and 75th percentiles with respect to both creative outcomes, specifically focusing on novelty and usefulness metrics.

By testing these hypotheses, we aim to identify the key factors that foster successful team formations and drive creative success in board game design.

4.2 Hyperedge and Descriptive Statistics Analysis

Variable	Value
Number of Boardgames	18965
Number of Designers	9857
Average Designer Gender (0=F,1=M)	0.79
Total Number of Teams	9072
#Teams of Size 1	5237
#Teams of Size 2	2719
#Teams of Size 3	765
#Teams of Size >3	351
Max Team Size	9
Mean Team Size	1.61
Number of Mechanics	192
Number of Mechanic Combinations	9608

Table 4.1: Cleaned BGG Dataset

The dataset after cleaning includes records of 18,965 game publications between 1990 and 2023. It comprises 9,857 designers (predominantly male, average gender score 0.79) and 9,072 unique teams. Most teams are single-designer (5,237), followed by two-member teams (2,719). Teams with three or more members are less common, with the maximum team size being nine and a mean size of 1.61. There are 192 distinct game mechanics, resulting in 9,608 unique mechanic combinations.

We use Event network analyzer (eventnet[1]) to further prepare the cleaned BGG data for analysis. Eventnet provides statistical analysis of networks of relational events by processing input files containing lists of events and computing explanatory variables (statistics) for all observations, including events and non-events. These generated non-events describe hyperedge statistics on events that could have happened at a given time point (but did not) and can then be used to estimate how covariates affect the relative event rate, i.e., the probability of an event happening.

To ensure computational tractability, we apply case-control sampling, where we sample 10 non-events for each observed event. We condition the risk sets (the set of all hyperedges on which an event could have taken place) to those hyperedges that contain the same number of designers and mechanics as the observed publication events. This approach allows us to model the factors influencing team formation and creative success without introducing significant bias or variation into our findings (for more detail see Lerner and Hâncean [2022]). This approach has the advantage of being computationally tractable but must be kept in mind during discussion. Essentially, it estimates which covariates make a collaboration more likely, assuming a fixed size of team members and a fixed set of mechanics.

We fit 3 models of increasingly complex sets of hyperedge statistics, resulting in parameter estimates for the respective hyperedge statistics. Significantly positive values indicate that higher values increase the probability of an event happening, while negative values suggest a decreased probability, and values around 0 imply no effect. These estimates will later be used for hypothesis testing and exploration. To keep an overview of the many statistics, we group them as follows:

- **RHEM Variables**: These describe first-order and repetition effects. For example, the gender ratio within a team or a measure of how many team members have previously collaborated.
- **RHOM Variables**: These include variables that reflect prior successes. For example, previous collaborative successes or success disparity within a team.
- **Mixed-Mode Variables**: These relate designers to their past usage of game mechanics. For example, how similar the game mechanics used by team members are to their previous projects.
- **Mixed-Mode Controls**: A set of statistics that will be included to control for main effects of mechanic usage. For example, base metrics like mechanic popularity and co-usage.

[1] Event network analyzer (eventnet): statistical analysis of networks of relational events. `https://github.com/juergenlerner/eventnet`

We will now examine the correlations of the main variables, i.e., the hyperedge statistics, as shown in Table 4.1.

Here are the key findings:

RHEM Effects

- **Team Size**:
  - Weak correlations with most variables.
  - Slight positive correlation with Team Sub Rep 3.
  - Slight positive correlation with Prior Success 2/3.
  - Slight positive correlation with Average Gender.
- **Gender Average and Gender Disparity**:
  - Generally low correlations with other variables, indicating minimal influence.
- **Tenure Average and Tenure Disparity**:
  - Tenure Average has a moderate correlation with Tenure Disparity.
  - Weak correlations with most other variables.
- **Team Closure**:
  - Moderate to high positive correlations with Team Exact Repetition (0.42) and Team Sub Rep 2/3 (0.62).
  - Positive correlation with Prior Success Exact (0.30).
- **Team Exact Repetition and Sub Repetitions**:
  - High positive correlation of Exact Repetition with Sub Rep 2 (0.93).
  - Moderate correlation with Team Sub Rep 1 (0.48).
  - Team Sub Rep 3 shows strong correlations with Team Closure (0.62).

RHOM Effects

- **Prior Success (Exact, 1, 2, 3, 4)**:
  - High correlation between Prior Success Exact and Prior Success 3 (0.85).
  - Prior Success 1 shows moderate correlation with Team Sub Rep 1 (0.65).
- **Performance Disparity**:
  - Generally weak correlations with other variables.

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
team_size	1		0	0.17	0.01	0.01	0.09	-0.03	0.02	0	0.16	0.14	0	0.04	0.04	0.15	0.13	0.06	0.02	0.02	0.01
gender_avg	2			0.07	0.04	0.02	0	0.01	0.06	0.01	-0.01	-0.02	-0.01	0.03	0	0	-0.01	0.05	0.01	-0.01	-0.01
gender_disparity	3				0.04	0.01	0.04	0.03	0.05	0.03	0.04	0.02	0.01	0.01	0.02	0.03	0.02	0.04	0.02	-0.02	-0.02
tenure_avg	4					0.5	0.05	0.04	0.27	0.03	0	0	-0.02	0	-0.02	-0.02	0	0.18	0.14	0.05	0.05
tenure_disparity	5						-0.09	-0.14	0.06	-0.16	-0.06	-0.03	-0.1	0.01	-0.11	-0.05	-0.01	0.1	0.04	0.06	0.05
team_closure	6							0.42	0.36	0.62	0.62	0.32	0.3	0.1	0.44	0.44	0.25	0.12	0.12	-0.03	-0.03
team_exact_repetition	7								0.48	0.93	0.27	0.08	0.58	0.07	0.52	0.21	0.05	0.07	0.19	-0.03	-0.03
team_sub_rep_1	8									0.51	0.14	0.06	0.26	0.24	0.31	0.1	0.05	0.64	0.35	-0.01	-0.02
team_sub_rep_2	9										0.35	0.14	0.55	0.09	0.6	0.27	0.1	0.09	0.2	-0.03	-0.04
team_sub_rep_3	10											0.4	0.35	0.09	0.36	0.79	0.29	0.02	0.05	-0.01	-0.02
team_sub_rep_4	11												0.09	0.03	0.12	0.25	0.67	0.03	0.03	-0.01	-0.01
prior_success_exact	12													0.21	0.85	0.43	0.12	0.08	0.11	-0.01	-0.02
prior_success_1	13														0.26	0.13	0.04	0.61	0.11	0	0
prior_success_2	14															0.46	0.17	0.14	0.13	-0.01	-0.02
prior_success_3	15																0.36	0.03	0.04	-0.01	-0.01
prior_success_4	16																	0.04	0.02	-0.01	-0.01
performance_disparity	17																		0.25	0.01	0.01
team_mechanic_closure	18																			0.46	0.56
portfolio_size_heterogeneity	19																				0.81
portfolio_size_avg	20																				

Table 4.1: Correlation Matrix of Main Variables

Mixed-Mode Effects

- **Team Mechanic Closure**:
  - Positive correlation with Portfolio Size Avg (0.54).
- **Portfolio Size Heterogeneity and Portfolio Size Average**:
  - Weak correlations with most variables.

Summary

The correlation matrix highlights various relationships within the groups of variables. The high correlation between Team Exact Repetition and Team Sub Rep 2 (0.93) indicates that pairs of team members often stick together. Male prevalence is suggested by the correlation between team size and gender average, where larger teams tend to have a higher ratio of men.

The RHEM variables show strong interrelations, particularly with team repetition patterns and tenure. RHOM variables, focusing on prior success and performance disparity, exhibit high correlations within prior success measures but remain largely independent of other variables. Mixed-Mode variables demonstrate correlations with other effects, suggesting that the mechanical aspects of team dynamics may influence broader performance metrics.

Correlations as high as 0.93 typically indicate a potential threat of collinearity among independent variables. We conducted VIF (Variance Inflation Factor) analyses on the models, which revealed values as high as approximately 27, implying severe multicollinearity. Since dropping the respective variables did not noticeably change the tendency of estimates and the standard errors of all models were less than 0.04, we decided to retain all variables.

4.3 Hyperedge Statistics and Collaboration Probability (H1)

The Cox Proportional Hazards (CoxPH) model was estimated using a stepwise approach to assess the impact of various factors on the probability of collaboration within designer teams. The procedure involved:

1. **Data Preparation**:
   - Only teams of size ≥ 2 were selected for the analysis.
   - All hyperedge statistics were standardized to ensure comparability.

2. **Conditioning**:
   - Models were conditioned on team size and the number of mechanics used.
   - This conditioning explains why estimates for team size and mechanic size are absent in the table.
   - For each event, 10 non-events were sampled.
   - The model assumes conditional independence, i.e. that events that occur at the same time are independent from one another.
3. **Model Fitting**:
   - Parameters were estimated using the 'coxph' function in the R-package 'survival'[2].
   - The RHEM model was fitted as the baseline to establish initial estimates.
   - For both Novelty and Usefulness:
     - The RHOM model was then fitted to incorporate prior successes.
     - The Designer x Mechanic model was fitted to include the effects of past usage of mechanisms.

Results Interpretation

The table of CoxPH coefficients (see 4.2) presents the impact of various factors on the probability of collaboration within designer teams. The distinction between novelty and usefulness as outcome variables refers to RHOM effects, which reflect past successes in terms of novelty or usefulness.

General Note on Interpretation: Positive coefficients indicate factors that increase the probability of collaboration, while negative coefficients suggest deterrents. Statistical significance is indicated as described in the table.

Estimation Procedure

1. **RHEM Model**: The base model coefficients provide a reference point for understanding the impact of fundamental variables on team formation regarding composition and past collaboration.
   - **Findings**:
     - A very small negative effect of gender average.

[2] https://CRAN.R-project.org/package=survival

	RHEM	RHOM		Designer x Mechanic	
	base	novelty	usefulness	novelty	usefulness
RHEM Effects					
team_size					
gender_avg	-0.05^{**}	-0.01	-0.05^{***}	-0.02	-0.05^{**}
gender_disparity	0.01	0.01	-0.01	-0.01	-0.02
tenure_avg	-0.72^{***}	-0.65^{***}	-0.66^{***}	-0.55^{***}	-0.54^{***}
tenure_disparity	-0.33^{***}	-0.40^{***}	-0.36^{***}	-0.26^{***}	-0.26^{***}
team_closure	-0.25^{***}	-0.19^{***}	-0.18^{***}	-0.16^{***}	-0.15^{***}
team_exact_repetition	-0.62^{***}	-0.50^{***}	-0.52^{***}	-0.30^{***}	-0.27^{***}
team_sub_rep_1	0.57^{***}	0.46^{***}	0.50^{***}	0.07	0.03
team_sub_rep_2	1.02^{***}	0.83^{***}	0.82^{***}	0.85^{***}	0.84^{***}
team_sub_rep_3	0.25^{***}	0.22^{***}	0.22^{***}	0.19^{***}	0.20^{***}
team_sub_rep_4	0.04^{**}	0.03^{+}	0.03	0.02^{+}	0.03^{*}
RHOM Effects					
prior_success_exact		0.08^{***}	0.02	0.01	-0.04^{**}
prior_success_1		-0.22^{***}	0.04	-0.10^{***}	0.12^{***}
prior_success_2		0.26^{***}	0.23^{***}	0.19^{***}	0.15^{***}
prior_success_3		-0.07^{***}	-0.08^{***}	-0.05^{***}	-0.03^{*}
prior_success_4		0.01	0.03^{*}	0.03	0.03^{+}
performance_disparity		0.10^{***}	0.00	0.04	0.03
Mixed-Mode Effects					
team_mechanic_closure				-0.07^{**}	-0.13^{***}
portfolio_size_heterogeneity				-0.00	0.00
portfolio_size_avg				0.15^{***}	0.18^{***}
Mixed-Mode Controls					
mechanic_size					
mechanics_exact_repetition				-0.12^{***}	-0.12^{***}
mechanic_sub_rep_1				0.60^{***}	0.61^{***}
mechanic_sub_rep_2				0.15^{***}	0.16^{***}
mechanic_sub_rep_3				0.01	0.01
mechanic_sub_rep_4				0.00	0.01
mechanic_sub_rep_5				0.00	0.01
exact_repetition				-0.01	-0.02^{**}
team_mechanic_sub_rep_1_1				0.46^{***}	0.48^{***}
team_mechanic_sub_rep_1_2				-0.07^{***}	-0.07^{***}
team_mechanic_sub_rep_1_3				0.01	0.01
team_mechanic_sub_rep_2_1				-0.42^{***}	-0.42^{***}
team_mechanic_sub_rep_2_2				0.03	0.04^{*}
team_mechanic_sub_rep_2_3				0.00	0.01
team_mechanic_sub_rep_3_1				-0.03	-0.05^{**}
team_mechanic_sub_rep_3_2				0.00	0.01
team_mechanic_sub_rep_3_3				-0.01^{+}	-0.02^{**}
AIC	26714.49	25648.61	26087.27	22805.11	22900.41
R^2	0.21	0.23	0.22	0.29	0.29

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$; $^{+}p < 0.1$

Table 4.2: The table presents the CoxPH coefficients for increasingly complex models. The base model with RHEM variables and RHOM and Mixed-Mode models for both novelty and usefulness as outcome variables.

  - A strong negative effect with respect to tenure average indicates that established developers are less likely to form teams.
  - A medium negative effect with respect to tenure disparity suggests that developers tend to work with others who have been around for a similar amount of time.
  - A medium negative effect of team closure indicates that teams that form usually do not have overlapping past collaborators.
  - A strong negative effect of team exact repetition suggests that groups are generally unlikely to repeat exact group efforts.
  - A strong positive effect of team sub-repetition 1 indicates that active developers are more likely to become part of a group.
  - Team sub-repetition 2 has the highest positive effect, meaning that groups of two are likely to collaborate more than once.
  - Team sub-repetition 3 and 4 also have positive effects, but these are less pronounced.

2. **RHOM Model (RHEM + RHOM)**: Incorporating prior successes, the model reveals how past achievements influence current team formation dynamics.

   - **Findings**:
     - Individual prior success with respect to novelty has a moderately negative effect on team formation.
     - Individual prior success with respect to usefulness has a moderately negative effect on team formation.
     - Prior success of size 2 with respect to novelty and usefulness has a moderately positive effect on team formation.
     - Adding RHOM effect lowered all RHEM coefficients slightly or kept them the same.
     - All other effects were negligible.

3. **Designer x Mechanic Model (RHEM + RHOM + Mixed-Mode)**: This model includes the past usage of mechanisms, providing insights into how prior interactions with specific mechanisms impact team formation.

   - **Findings**:
     - For novelty, there is a very small negative team-mechanic-closure effect, indicating that team members typically collaborate with those who have used different mechanics in the past. This effect is slightly stronger with respect to usefulness.
     - Experience with many mechanics is positively related to group formation, with a similar effect for usefulness.
     - Estimates for RHEM and RHOM effects changed slightly but remained consistent in tendency.

Conclusion and Model Selection

The analysis reveals that factors such as tenure, gender, past collaboration, and prior successes significantly impact team formation, with dominant negative effects from tenure, light negative team closure and strong positive effects from repeated sub-teams. All models have low standard errors (<0.05), indicating reliable fitting despite potential collinearity of independent variables, likely due to the conditioning on team and mechanic size.

Among the models, the Designer x Mechanic Model (RHEM + RHOM + Mixed-Mode) exhibits the best performance, shown by superior AIC and R^2 metrics. The Akaike Information Criterion (AIC) is a measure of the relative quality of statistical models for a given set of data; a lower AIC indicates a better-fitting model. The R^2 metric measures the proportion of variance in the dependent variable predictable from the independent variables–higher values indicate a better fit. Therefore, this model is selected for further investigation in H2 and H3, as it provides the most comprehensive insights into the factors influencing collaboration probability within designer teams.

4.4 Correlation between Collaboration Probability and Creative Outcomes (H2)

To investigate the relationship between the factors driving collaboration and creative outcomes, we performed an OLS regression on the covariates used in the Cox Proportional Hazards (CoxPH) model. This method, as described in Lerner and Hâncean [2022], involves fitting a Generalized Linear Model (GLM) on Hyperedge statistics. While CoxPH models the relative event rate, conceptualized here as collaboration probability, using both events and sampled non-events, the OLS regression focuses solely on events, modeling the conditional distribution of outcomes based on specific hyperedge statistics. CoxPH coefficients identify drivers of collaboration, whereas OLS coefficients highlight drivers of success.

Given the complexity of interpreting the table of variables, we plotted the CoxPH coefficients against the OLS coefficients for both novelty and usefulness, as shown in Figure 4.3. The x-axis represents the magnitude of CoxPH coefficients, indicating whether a covariate increases or decreases collaboration probability. The y-axis represents the magnitude of OLS coefficients, showing the relationship between statistics and creative outcomes. Creative outcomes are marked by triangles (novelty) and circles (usefulness), connected by lines for convenience. Colors indicate groups of variables (RHEM, RHOM, mixed mode[3]). Only significant estimates are displayed.

[3] Controls were included for insights on mechanic usage behavior, despite their indirect relation to team formation

Interpretation: The primary question is whether the covariate estimates move in the same direction. This can be assessed by their placement in the plot quadrants:

- **Q1 (top left):** Low collaboration, high success
- **Q2 (top right):** High collaboration, high success
- **Q3 (bottom left):** Low collaboration, low success
- **Q4 (bottom right):** High collaboration, low success

In terms of adverse selection, i.e., whether certain factors might lead to increased collaboration but do not necessarily lead to successful outcomes, or vice versa. This misalignment can indicate suboptimal team selection in the board game industry.

- **Q1 and Q4 indicate adverse selection**, where covariates lead to either low collaboration with high success (Q1) or high collaboration with low success (Q4). This suggests that the factors driving collaboration may not always align with those driving creative success, leading to situations where collaboration occurs without resulting in successful creative outcomes, or successful outcomes occur without much collaboration.
- **Q2 and Q3 indicate non-adverse selection**, where covariates lead to high collaboration with high success (Q2) or low collaboration with low success (Q3). This suggests that the factors influencing collaboration and creative success are aligned, resulting in efficient and effective team performance.

Findings:

1. **RHEM Effects**
   (a) **Team closure** in Q1 for novelty: Teams of members connected by common past collaborators are less likely to collaborate but, if they do, they typically create slightly more novel games.
   (b) **Team exact repetition** in Q2 for novelty: Teams are less likely to join again with the exact team, and if they do, it typically results in less novel games.
   (c) **Team subrep3 and team subrep2** for usefulness in Q4: Teams that have previously collaborated publish games which receive slightly worse ratings than others.
   (d) Overall, RHEM effects were relatively high in CoxPH coefficients and low in OLS coefficients.
2. **RHOM Effects**

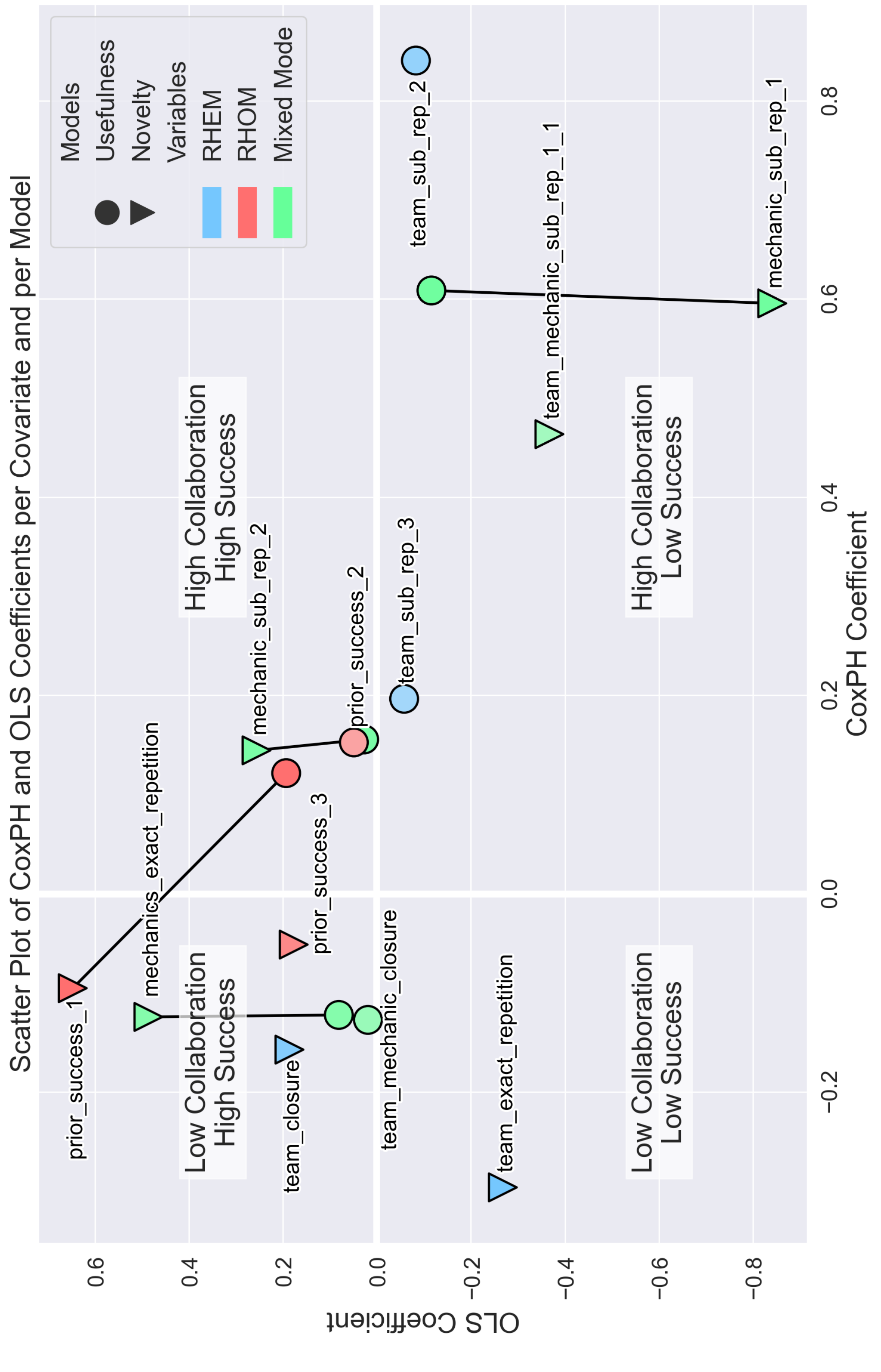

Figure 4.3: Correlation between Collaboration Probability and Creative Outcomes

(a) **Prior success 3** in Q1 for novelty: Teams of 3 that previously published novel games are unlikely to form, but if they do, they often achieve more novel games.

(b) **Prior success 2** in Q2 for usefulness: Teams of 2 that previously published successful games are slightly more likely to publish better-rated games.

(c) **Prior success 1**: For novelty in Q1 and for usefulness in Q2. This means that innovators are less likely to participate in teams but, if they do, are very likely to innovate again. At the same time, successful individuals with respect to usefulness are more likely to participate in joint game development, which typically receives better ratings. This is one of the stronger and more interesting effects.

3. **Mixed Mode Effects**

   (a) **Team mechanic closure** in Q1 for usefulness: Teams that have used different mechanics in the past tend to create games which receive better ratings. However, this effect is almost negligible.

   (b) **Team mechanic sub rep 1 1** in Q4 for novelty: This effect is relatively large and indicates that team members who frequently use the same mechanic are likely to develop less innovative games.

4. **Mixed Mode Controls**

   (a) Note: This does not directly describe team formation and its consequences but provides insights into mechanic usage within the BGG community.

   (b) **Mechanics exact repetition**

      i. Q1 for both novelty (stronger effect) and usefulness.

   (c) **Mechanics subrep1**

      i. Q4 for both novelty (stronger negative effect) and usefulness.

   (d) **Mechanics subrep2**

      i. Q3 for both novelty and usefulness.

   (e) Usefulness effects are rather small, but novelty effects are large and, at first glance, seem to contradict the repetition statistics of different sizes:

      i. Using always the same mechanic is extremely negatively related to novelty.

      ii. Prevalence of using 2 in combination or exactly repeating mechanics is positively related to novelty.

   (f) **Methodological Note:** Interpreted jointly, it seems like subrep1 overestimates the negative effect; the other two related statistics balance this by the positive estimates. This effect with alternating sign of subset repetition has been described as a potential weakness by Lerner and Hâncean [2022] of the RHEM/RHOM models but will probably be addressed in the future. Accordingly, the authors recommend joint interpretation of (subset) repetitions of different sizes.

Summary H2:

Alignment between Collaboration and Success Teams with prior successful game publications are more likely to achieve high ratings if they collaborate again (2 b). Innovators tend to avoid team participation but are likely to innovate when they do join, while successful individuals frequently contribute to high-rating games (2 c). Teams using varied mechanics receive better ratings, though this effect is minimal (3 a). Combining or repeating two mechanics positively impacts novelty, suggesting certain patterns of repetition and prior success can influence the likelihood of producing innovative and successful games (4 d).

Misalignment between Collaboration and Success Teams with members who have collaborated with common past collaborators tend to create slightly more novel games, despite being less likely to collaborate again (1 a). Teams that repeat exactly the same composition tend to produce less novel games (1 b). Previous collaboration generally leads to slightly worse ratings for games, indicating a possible misalignment between factors driving collaboration and creative success (1 c). Additionally, frequent use of the same mechanic by team members is linked to less innovative games, and using the same mechanics repeatedly is detrimental to novelty (3 b and 4 a).

4.5 Comparison between High-Creativity and Low-Creativity Teams (H3)

In the final part of the thesis, we investigate whether high-creativity teams form differently from low-creativity teams. We compare formation dynamics of the 25th and 75th percentiles with respect to both creative outcomes.

This analysis focuses on the standardized hyperedge statistics of the respective groups rather than the fitted coefficients. As comparisons using simple scores in tables can be difficult to interpret, we included a heatmap-style coloring in the figure 4.4. Rows represent the hyperedge statistics, and columns represent the groups at the 25th and 75th percentiles. The split between novelty and usefulness is due to the RHOM variables reflecting prior success concerning the respective measures.

We conducted Shapiro-Wilk normality tests on all groups and hyperedge statistics. Since the data are not normally distributed, we used the Mann-Whitney U test to assess the significance of differences between the groups. If the difference between the groups is significant ($p<0.05$), the field is colored according to the heatmap.

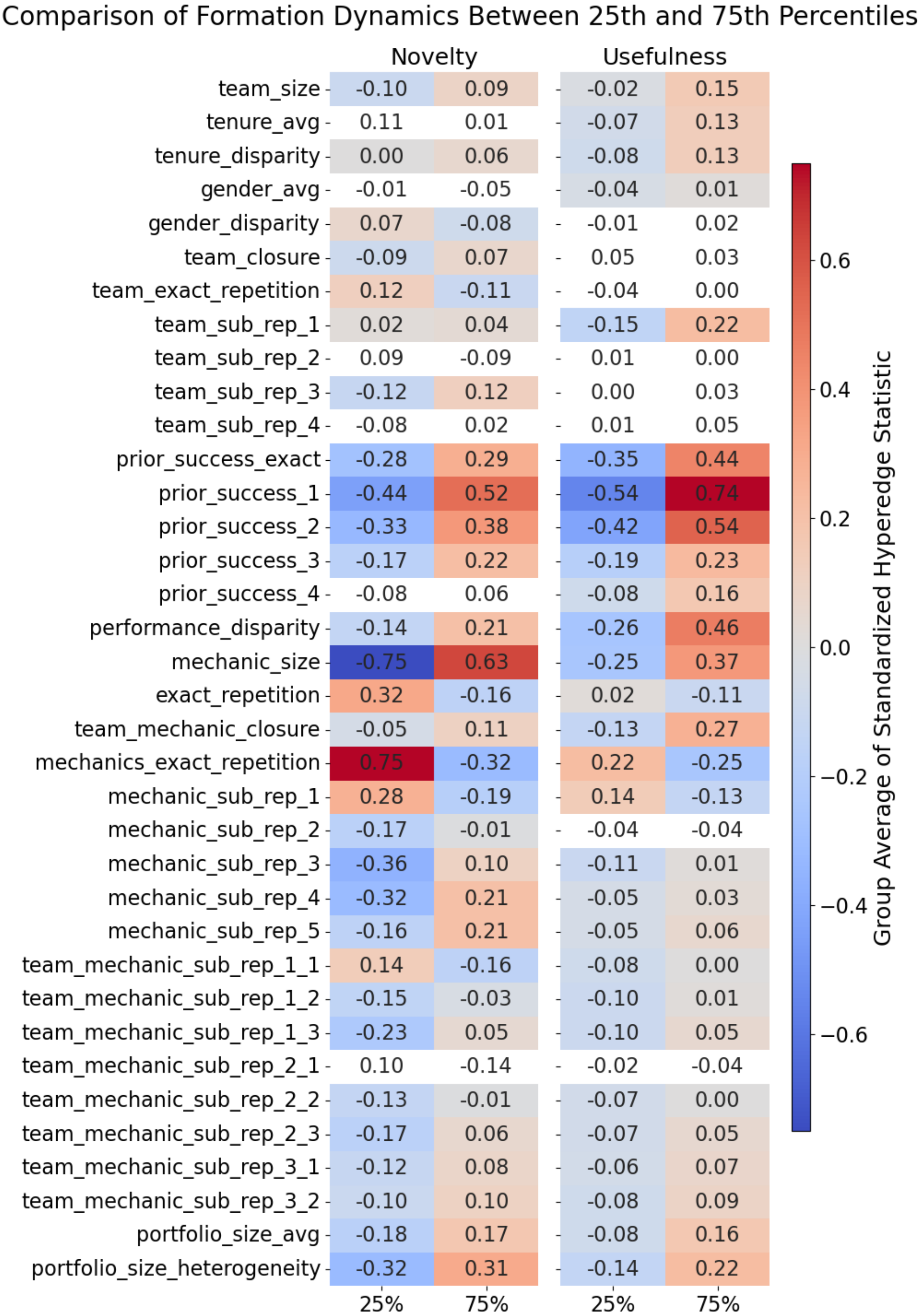

Figure 4.4: Heatmap comparison of standardized hyperedge statistics between high-creativity and low-creativity teams, highlighting significant differences in formation dynamics at the 25th and 75th percentiles (of respective creative outcomes).

Findings:

1. **RHEM Variables**
   (a) Overall, RHEM differences were small (≤ 0.37).
   (b) In the novelty-reflecting model, the high-performing teams were:
      i. Slightly larger
      ii. Had more tenure and gender disparity
      iii. Exhibited more team closure
      iv. Had slightly more experience working together
   (c) In the usefulness-reflecting model, the high-performing teams were:
      i. Slightly larger
      ii. Showed higher tenure and higher tenure disparity
      iii. Gender differences were negligible
      iv. Had slightly more experience working together
      v. Individuals in the group were more active (strong RHEM difference)
2. **RHOM Variables**
   (a) RHOM differences were the strongest (≤ 1.28).
   (b) In both the novelty and usefulness reflecting models, the high-performing teams:
      i. Had more past successes across the measures (strongest effect for prior success of order 1, with slightly stronger effects for usefulness than for novelty)
      ii. Had higher within-team success disparities
3. **Mixed Mode Variables**
   (a) In the novelty-reflecting model, the high-performing teams:
      i. Used more mechanics (size) – very strong effect
      ii. Rarely used the same combination of mechanics
      iii. Typically had team members with different portfolio sizes
   (b) In the usefulness-reflecting model, the high-performing teams:
      i. Used more mechanics (size)
      ii. Had, on average, experimented with/used more distinct mechanics in the past (portfolio size average)
      iii. Had more experience with the same mechanics (team mechanic closure)
      iv. Else negligible differences

The analysis revealed that high-creativity teams, compared to low-creativity teams, differ significantly in several areas. For RHEM variables, high-creativity teams were generally larger (1.a.i), had more tenure and gender disparity (1.a.ii), exhibited more team closure (1.a.iii), and had slightly more experience working together (1.a.iv). RHOM variables showed the strongest differences, with high-creativity teams having more past successes (2.a.i) and higher within-team success disparities (2.a.ii). For Mixed Mode variables, high-creativity teams used a greater variety of mechanics (3.a.i), rarely used the same combination of mechanics (3.a.ii), and typically had team members with different portfolio sizes (3.a.iii). High-creativity teams also had more experience with the same mechanics (3.b.iii), though other differences were negligible.

4.6 Summary of Findings

- The main drivers of team formation (RQ1) were identified through hyperedge statistics using Cox Proportional Hazards models. Key factors included tenure, gender ratio, past collaboration patterns, and prior successes. Teams are more likely to form when they have members with similar tenure and prior collaboration experiences, with repeated sub-team collaborations being particularly significant.
- The relationship between team formation drivers and creative success (RQ2) was explored using OLS regression. Findings indicated both alignment and misalignment between collaboration probability and creative outcomes. Teams with prior success and varied mechanical experience tended to achieve higher creative success, while repeated exact team compositions and frequent use of the same mechanics correlated with lower novelty.
- High-creativity teams (H3) differed significantly from low-creativity teams in several aspects. They were generally larger, had greater tenure and slight gender disparities, exhibited more team closure, and had most of all more past successes. High-creativity teams also demonstrated more variety and experience with different mechanics.
- The analysis revealed that factors such as team size, past collaboration, and mechanic usage significantly impact both the formation and success of design teams, contributing to a comprehensive understanding of how network effects influence creative outcomes in board game design.

Chapter 5

Discussion

Interpretation of Results

The results of this study provide valuable insights into the dynamics of team formation and creative success within the board game industry. By analyzing the factors that influence collaboration and their relationship to creative outcomes, we can better understand the mechanisms driving innovation in self-organized teams.

Key Drivers of Team Formation

Our analysis identified several key factors that significantly impact the likelihood of team formation. Repetition effects, such as previous collaborations among team members, were strong predictors of future collaborations, aligning with social network theory, which suggests that established relationships and trust facilitate repeated interactions. Additionally, tenure disparity was negatively influential, indicating that teams with members having similar tenure lengths were more likely to form. Teams with members who have low tenure were common, suggesting that a balance between diversity and familiarity is crucial for successful collaboration. However, the effects of team demographic diversity and disparity were generally small to negligible, highlighting that other factors may play more pivotal roles in team formation.

Relationship Between Team Formation and Creative Success

The relationship between collaboration factors and creative success was complex. Some factors, such as prior individual success and diverse mechanical experience, were positively correlated with both collaboration likelihood and creative outcomes. However, other factors showed misalignment. For instance, teams frequently using the same mechanics or having high exact repetition were less likely to produce novel games, despite their tendency to form. This indicates that while familiarity and repeated interactions foster collaboration, they may limit creativity if not balanced with fresh perspectives and innovative approaches.

High-Creativity vs. Low-Creativity Teams

High-creativity teams differed significantly from low-creativity teams. They tended to be larger, with greater tenure and gender diversity, and possessed more varied mechanical experience. These teams also showed higher levels of team closure and prior success, suggesting that a combination of diverse skills, established relationships, and past achievements contributes to high creative performance. This highlights the importance of fostering environments that support both diversity and collaboration to drive innovation.

Parallels with Organizational Psychology and Network Theory

In organizational psychology, the understanding of team creativity has evolved from linear Input-Process-Output (IPO) models to more recursive and complex Input-Mediator-Output-Input (IMOI) models. This shift acknowledges that team processes are dynamic and cyclic rather than straightforward and linear. Similarly, our study reflects this complexity by examining how network effects and team dynamics influence creative outcomes over time.

Although we do not directly observe psychological processes in this study, our approach of relating network effects with creative outcomes parallels the IMOI model. By considering factors such as past collaborations, tenure, and diversity, we capture the recursive nature of team dynamics where past experiences influence current interactions, which in turn affect future creative performance. This perspective aligns with the recursive feedback loops emphasized in IMOI models, where inputs (e.g., team composition) and mediators (e.g., collaboration patterns) interact over time to influence outcomes (e.g., creativity).

Network Theory and Relational Events

Our study also connects with network theory by focusing on relational events, which are crucial for understanding social processes and structures. Relational event models allow us to create realistic null models and delve deeply into social interactions, offering a nuanced view of how team dynamics unfold over time. By examining homophily—the tendency for similar individuals to associate with one another—we utilize a framework that captures these dynamics within realistic contexts.

While homophily was not explicitly a focal point in our results, the tendencies observed in our analysis suggest its presence. The significant role of prior collaborations and tenure in team formation indicates a preference for working with familiar individuals, aligning with homophily. This tendency highlights the importance of balancing familiarity and diversity, as excessive homophily can stifle creativity by limiting diversity, while too much diversity without cohesion can hinder effective collaboration.

Team Closure and Subset Repetition

We found negative team closure and positive subset repetition. Subset repetition, when controlling for exact repetition, indicates dense clusters in the network of designers. Together with negative team closure, this suggests the presence of communities within the network that collaborate but do not merge over time. Interpreted in terms of access to knowledge and skills, this allows for creative variety, as different methods of work within subgroups can evolve, potentially engaging distinct communities of players and users. The observation that the 75th percentile of teams has more negative closure compared to the 25th percentile suggests that innovative teams leverage knowledge from separate groups, finding ways to combine diverse people and their expertise.

Prior Success

Prior success may indicate skill, but it may also reflect preferential attachment or "rich-get-richer" effects. This is especially likely regarding prior success and usefulness, where fans might rate a game more favorably than a random game. In terms of novelty as a measure of success, it is more likely to indicate a capacity to innovate. Innovative sub-teams in the past are less likely to collaborate again, but when they do, they tend to innovate more.

Methodological Considerations

In this study, we employed the relatively new Relational Hyper Event Model (RHEM) and Relational Hyper Event Outcome Model (RHOM), as proposed by [Lerner and Hâncean, 2022]. These models provide a general statistical framework for explaining collaboration rates and the determinants of creative outcomes in complex networks. While these methods are powerful, they are also complex and challenging to work with and interpret.

We see significant potential in the application of RHEM and RHOM to practical contexts like the board game industry. However, the exploratory nature of our approach posed some challenges. Ideally, future research should have specific network effects in mind, formulate alternative hypotheses, and rigorously test these hypotheses. This focused approach would likely mitigate issues related to collinearity among independent variables—a topic that we acknowledge was not adequately addressed in this study.

Limitations and Future Research

Despite its contributions, this study has several limitations.

One obvious limitation is the reliance on observational data, which restricts our ability to draw causal conclusions. This is inherent to this kind of research. Beyond that, there are several methodological limitations that should be addressed in future studies:

First, a more thorough approach to address collinearity is necessary. Finding principled ways to deal with this issue will improve the robustness of the findings.

Second, the conditional size model simplifies the comparison of highly size-dependent network statistics. Future work should aim to fit the hyperedge size in expectation for sampled non-events, which would allow for a more accurate assessment of the size effect—an important factor in team design as indicated by meta-studies.

Third, we assumed that the effect of past collaborations (all hyperedge statistics) does not decrease over time. Although the models allow for the inclusion of decay effects, we chose not to incorporate this due to the existing complexity. Future studies should include time decay, motivated by theory, to reflect more realistic network effects.

Fourth, while the models can test interactions between covariates and curvilinear relationships, we opted for simplicity given our research questions. Future research should explore these aspects to provide a more nuanced understanding of the dynamics at play.

Fifth, novelty, as an outcome variable, is unique as it encodes the past usage of mechanics. By including statistics on designer-mechanic interactions, we are essentially predicting a global complex network statistic (novelty) using more localized statistics. The theoretical and practical implications of this approach, along with potential endogeneity issues, should be addressed in future studies.

Sixth, our dataset records every rating for every game individually, but we did not investigate the community aspect. This could be a rich source for new analyses, enabling questions about how the audience responds to innovation and how audience and creators co-evolve and co-create. This is particularly relevant in creativity research, as creativity always reflects its time and context and ultimately relies on an audience.

Lastly, due to the limited scope of this thesis, we ignored game-level covariates such as complexity, category, and many other specifics. Including these variables would have made the model more complicated, but they should be included in future research for a more comprehensive quantitative analysis.

While this study focused on the board game industry, the findings (that collaboration captured as hyperedges serves as an explanatory variable) probably generalize to other creative sectors. We lack a practical question and theoretical narrow lens, but it would be interesting to test a theory using RHEM/RHOM in different creative industries. Comparative studies across various industries could investigate general patterns in collaborative creativity, providing a more comprehensive understanding of the factors driving team creativity.

Chapter 6

Conclusion

This study provides a comprehensive analysis of the dynamics of team formation and creative success in the board game industry. By leveraging network effects and applying the novel Relational Hyper Event Model (RHEM) and Relational Hyper Event Outcome Model (RHOM), several key factors influencing team formation and creative outcomes were identified. These models offer a nuanced understanding of the mechanisms driving innovation in self-organized teams.

The analysis revealed that repetition effects and tenure similarity are significant drivers of team formation. Established relationships and trust were strong predictors of future collaborations, consistent with social network theory. While demographic diversity had minimal impact on team formation, tenure similarity played a crucial role, suggesting that familiarity fosters collaboration.
The relationship between collaboration drivers and creative success was complex. Prior individual success and diverse mechanical experience were positively correlated with both collaboration likelihood and creative outcomes. However, excessive familiarity and repeated interactions without innovative approaches were found to limit creativity. High-creativity teams differed significantly from low-creativity teams. They were generally larger, with greater tenure and gender diversity, and possessed more varied mechanical experience. These teams also exhibited higher levels of team closure and prior success, indicating that a combination of diverse skills, established relationships, and past achievements contributes to high creative performance.
The findings underscore the importance of balancing familiarity and diversity in fostering creative collaborations. Encouraging diverse mechanical experiences and leveraging prior successes while avoiding excessive repetition of team compositions can enhance creative outcomes. These insights can inform strategies for fostering innovation in the board game industry and other creative sectors.

Future research should address the methodological limitations identified in this study. Specifically, addressing collinearity among independent variables, incorporating time decay effects, and exploring interactions and curvilinear relationships between covariates will enhance the robustness and accuracy of the findings. Additionally, including game-level specifics and investigating the community aspect of ratings will provide a more comprehensive understanding of the factors driving team creativity.
Comparative studies across different creative industries could also provide broader insights into the generalizability of the findings. By testing the applicability of RHEM and RHOM in various contexts, future research can further elucidate the dynamics of collaborative creativity and its impact on innovation.

In conclusion, this study advances our understanding of the dynamics of team formation and creative success in the board game industry. By identifying the key drivers of collaboration and their relationship to creative outcomes, and through the application of RHEM and RHOM, it provides a foundation for researching innovation in self-organized teams.

Appendix A

Code Repository

All the code used in this thesis is available at the following repository:

- GitHub: `https://github.com/creamartin/master_css`

Instructions

To clone the repository, use the following command:

```
git clone https://github.com/creamartin/master_css
```

For detailed instructions on how to run the code, please refer to the README file in the repository.

Appendix B

Supplementary Information

Category	Details
Game Identification	Unique identifiers for each game. Titles of the games.
Publication Information	Year of release.
Game Type	Categorization of games into different types.
Contributors	Information about the designers, artists, and publishers of the games.
Audience/Player-related Information	Minimum and maximum number of players. Recommended and optimal number of players for best gameplay experience. Age appropriateness – Minimum and recommended age for players. Playtime – Minimum and maximum duration of gameplay.
Game Characteristics	Categories and mechanics of the games. Indication of whether a game is cooperative. Compilation status and details about specific game compilations. Family and implementation information.
Popularity and Engagement	Popularity ranking on BGG. Number of votes received from users. Average user ratings and Bayesian average ratings. Standard deviation of ratings indicating consistency.
Complexity and Accessibility	Complexity rating indicating the difficulty level of the game. Language dependency showing the necessity of language comprehension for gameplay.
Game Implementations	Different versions or editions of the game.

Table B.1: Detailed Information at the Board Game Level

VIFS Scores

Variable	VIF	Most_Correlated_Var	Correlation
team_sub_rep_2	15.27	team_exact_repetition	0.89
team_exact_repetition	11.31	team_sub_rep_1	0.45
team_mechanic_sub_rep_2_1	8.87	team_sub_rep_2	0.75
team_sub_rep_3	8.58	team_mechanic_sub_rep_2_3	0.11
team_mechanic_sub_rep_3_1	8.01	team_sub_rep_3	0.87
team_mechanic_sub_rep_1_1	7.62	team_mechanic_sub_rep_1_3	0.33
team_sub_rep_1	7.28	tenure_avg	0.69
team_mechanic_sub_rep_1_2	6.77	team_mechanic_sub_rep_2_1	0.47
team_mechanic_sub_rep_2_2	6.56	team_mechanic_sub_rep_1_2	0.77
team_mechanic_sub_rep_1_3	6.29	team_mechanic_sub_rep_2_2	0.50
team_mechanic_sub_rep_2_3	5.93	team_mechanic_sub_rep_1_3	0.84
prior_success_2	5.40	prior_success_exact	0.79
prior_success_exact	4.46	prior_success_1	0.54
team_mechanic_sub_rep_3_2	4.34	team_mechanic_sub_rep_3_1	0.80
mechanic_sub_rep_2	4.21	mechanic_sub_rep_1	0.61
portfolio_size_avg	3.84	portfolio_size_heterogeneity	0.81
mechanic_sub_rep_1	3.71	mechanic_sub_rep_1	1
portfolio_size_heterogeneity	3.67	portfolio_size_heterogeneity	1
prior_success_3	3.60	team_sub_rep_3	0.77
performance_disparity	3.59	team_sub_rep_1	0.63
mechanics_exact_repetition	3.49	mechanic_size	−0.61
team_closure	3.46	team_sub_rep_2	0.48
prior_success_1	3.07	prior_success_1	1
mechanic_size	2.80	mechanic_size	1
mechanic_sub_rep_3	2.61	mechanic_sub_rep_3	1
team_sub_rep_4	2.40	prior_success_3	0.22
prior_success_4	2.26	team_sub_rep_4	0.67
team_mechanic_sub_rep_3_3	2.25	team_mechanic_sub_rep_3_2	0.65
team_mechanic_closure	2.22	portfolio_size_heterogeneity	0.46
team_size	2.21	team_sub_rep_2	0.02
mechanic_sub_rep_4	2.18	mechanic_sub_rep_4	1
tenure_avg	2.14	team_exact_repetition	0.31
mechanic_sub_rep_5	1.58	mechanic_sub_rep_4	0.59
tenure_disparity	1.56	prior_success_4	0.02
exact_repetition	1.54	team_mechanic_sub_rep_2_1	0.50
gender_disparity	1.10	gender_avg	0.20
gender_avg	1.09	gender_avg	1

Table B.1: VIF Scores for OLS on Novelty.

Variable	VIF	Most_Correlated_Var	Correlation
team_sub_rep_2	15.27	team_exact_repetition	0.89
team_exact_repetition	11.48	team_sub_rep_1	0.45
team_mechanic_sub_rep_2_1	8.89	team_sub_rep_2	0.75
team_sub_rep_1	8.36	tenure_avg	0.69
team_mechanic_sub_rep_3_1	8.09	team_sub_rep_3	0.87
team_mechanic_sub_rep_1_1	7.66	team_mechanic_sub_rep_1_3	0.33
team_sub_rep_3	6.99	team_mechanic_sub_rep_2_3	0.11
team_mechanic_sub_rep_1_2	6.79	team_mechanic_sub_rep_2_1	0.47
team_mechanic_sub_rep_2_2	6.56	team_mechanic_sub_rep_1_2	0.77
team_mechanic_sub_rep_1_3	6.29	team_mechanic_sub_rep_2_2	0.50
team_mechanic_sub_rep_2_3	5.91	team_mechanic_sub_rep_1_3	0.84
prior_success_2	5.49	prior_success_exact	0.81
prior_success_exact	4.73	prior_success_1	0.57
team_mechanic_sub_rep_3_2	4.35	team_mechanic_sub_rep_3_1	0.80
performance_disparity	4.32	team_sub_rep_1	0.67
mechanic_sub_rep_2	4.24	mechanic_sub_rep_1	0.61
portfolio_size_avg	3.84	portfolio_size_heterogeneity	0.81
mechanic_sub_rep_1	3.71	mechanic_sub_rep_1	1
portfolio_size_heterogeneity	3.67	portfolio_size_heterogeneity	1
mechanics_exact_repetition	3.50	mechanic_size	−0.61
team_closure	3.44	team_sub_rep_2	0.48
prior_success_1	3.40	prior_success_1	1
mechanic_size	2.78	mechanic_size	1
prior_success_3	2.77	team_sub_rep_3	0.66
mechanic_sub_rep_3	2.61	mechanic_sub_rep_3	1
team_mechanic_closure	2.28	portfolio_size_heterogeneity	0.46
team_mechanic_sub_rep_3_3	2.26	team_mechanic_sub_rep_3_2	0.65
team_size	2.25	team_sub_rep_2	0.02
mechanic_sub_rep_4	2.18	mechanic_sub_rep_4	1
tenure_avg	2.16	team_exact_repetition	0.31
team_sub_rep_4	1.98	prior_success_3	0.21
prior_success_4	1.90	team_sub_rep_4	0.56
mechanic_sub_rep_5	1.58	mechanic_sub_rep_4	0.59
tenure_disparity	1.54	prior_success_4	0.05
exact_repetition	1.50	team_mechanic_sub_rep_2_1	0.50
gender_disparity	1.10	gender_avg	0.20
gender_avg	1.09	gender_avg	1

Table B.2: VIF Scores for OLS on Usefulness.

Variable	VIF	Most_Correlated_Var	Correlation
team_sub_rep_2	27.00	team_exact_repetition	0.93
team_sub_rep_1	16.48	performance_disparity	0.64
team_exact_repetition	16.48	team_sub_rep_2	0.93
team_sub_rep_3	12.31	team_mechanic_sub_rep_3_1	0.87
team_mechanic_sub_rep_1_1	11.85	team_mechanic_sub_rep_2_1	0.75
team_mechanic_sub_rep_2_1	10.89	team_sub_rep_2	0.85
team_mechanic_sub_rep_3_1	10.05	team_sub_rep_3	0.87
team_mechanic_sub_rep_2_2	6.77	team_mechanic_sub_rep_1_2	0.81
team_mechanic_sub_rep_1_2	6.65	team_mechanic_sub_rep_2_2	0.81
team_mechanic_sub_rep_1_3	5.66	team_mechanic_sub_rep_2_3	0.85
prior_success_2	5.49	prior_success_exact	0.85
team_mechanic_sub_rep_2_3	5.38	team_mechanic_sub_rep_1_3	0.85
mechanic_sub_rep_2	4.73	mechanic_sub_rep_1	0.64
prior_success_exact	4.66	prior_success_2	0.85
team_mechanic_sub_rep_3_2	4.62	team_mechanic_sub_rep_3_1	0.80
mechanic_sub_rep_1	4.41	mechanic_sub_rep_2	0.64
team_closure	4.33	team_sub_rep_3	0.62
prior_success_3	3.99	team_sub_rep_3	0.79
portfolio_size_heterogeneity	3.70	portfolio_size_avg	0.81
portfolio_size_avg	3.62	portfolio_size_heterogeneity	0.81
mechanics_exact_repetition	3.56	mechanic_sub_rep_1	0.43
performance_disparity	3.37	team_sub_rep_1	0.64
mechanic_sub_rep_3	2.71	mechanic_sub_rep_2	0.62
prior_success_1	2.66	performance_disparity	0.61
tenure_avg	2.63	tenure_disparity	0.50
team_mechanic_closure	2.55	portfolio_size_avg	0.56
team_sub_rep_4	2.46	prior_success_4	0.67
team_mechanic_sub_rep_3_3	2.42	team_mechanic_sub_rep_3_2	0.65
mechanic_size	2.34	portfolio_size_heterogeneity	0.40
mechanic_sub_rep_4	2.21	mechanic_sub_rep_3	0.61
team_size	2.11	gender_disparity	0.17
prior_success_4	2.10	team_sub_rep_4	0.67
exact_repetition	1.69	team_mechanic_sub_rep_2_1	0.50
mechanic_sub_rep_5	1.65	mechanic_sub_rep_4	0.60
tenure_disparity	1.47	tenure_avg	0.50
gender_disparity	1.10	team_size	0.17
gender_avg	1.09	gender_disparity	0.07

Table B.3: VIF Scores for CoxPH on Novelty.

Variable	VIF	Most_Correlated_Var	Correlation
team_sub_rep_2	28.71	team_exact_repetition	0.93
team_sub_rep_1	19.20	performance_disparity	0.71
team_exact_repetition	18.19	team_sub_rep_2	0.93
team_mechanic_sub_rep_1_1	12.46	team_mechanic_sub_rep_2_1	0.75
team_mechanic_sub_rep_2_1	11.51	team_sub_rep_2	0.85
team_mechanic_sub_rep_3_1	10.35	team_sub_rep_3	0.87
team_sub_rep_3	10.15	team_mechanic_sub_rep_3_1	0.87
team_mechanic_sub_rep_2_2	6.86	team_mechanic_sub_rep_1_2	0.81
team_mechanic_sub_rep_1_2	6.74	team_mechanic_sub_rep_2_2	0.81
team_mechanic_sub_rep_1_3	6.09	team_mechanic_sub_rep_2_3	0.85
team_mechanic_sub_rep_2_3	5.78	team_mechanic_sub_rep_1_3	0.85
prior_success_2	5.18	prior_success_exact	0.86
mechanic_sub_rep_2	4.79	mechanic_sub_rep_1	0.65
prior_success_exact	4.78	prior_success_2	0.86
team_mechanic_sub_rep_3_2	4.68	team_mechanic_sub_rep_3_1	0.80
mechanic_sub_rep_1	4.49	mechanic_sub_rep_2	0.65
team_closure	4.28	team_sub_rep_3	0.62
performance_disparity	4.19	team_sub_rep_1	0.71
portfolio_size_heterogeneity	3.55	portfolio_size_avg	0.81
mechanics_exact_repetition	3.48	mechanic_sub_rep_1	0.43
portfolio_size_avg	3.38	portfolio_size_heterogeneity	0.81
prior_success_3	3.34	team_sub_rep_3	0.68
prior_success_1	3.29	performance_disparity	0.70
team_mechanic_closure	2.76	portfolio_size_avg	0.54
mechanic_sub_rep_3	2.72	mechanic_sub_rep_2	0.62
team_sub_rep_4	2.60	prior_success_4	0.57
tenure_avg	2.60	tenure_disparity	0.50
team_mechanic_sub_rep_3_3	2.51	team_mechanic_sub_rep_3_2	0.65
prior_success_4	2.30	team_sub_rep_4	0.57
mechanic_size	2.30	portfolio_size_heterogeneity	0.40
mechanic_sub_rep_4	2.19	mechanic_sub_rep_3	0.61
team_size	2.10	gender_disparity	0.17
mechanic_sub_rep_5	1.65	mechanic_sub_rep_4	0.60
exact_repetition	1.64	team_mechanic_sub_rep_2_1	0.50
tenure_disparity	1.47	tenure_avg	0.50
gender_disparity	1.10	team_size	0.17
gender_avg	1.08	team_sub_rep_1	0.06

Table B.4: VIF Scores for CoxPH on Usefulness.

Bibliography

[1] Amabile, T. (1983). The social psychology of creativity: A componential conceptualization. *Journal of personality and social psychology*, 45(2):357.

[2] Amabile, T. M. and Pratt, M. G. (2016). The dynamic componential model of creativity and innovation in organizations: Making progress, making meaning. *Research in Organizational Behavior*, 36:157–183.

[3] Baten, R. A. (2021). Fantastic ideas and where to find them: Elevating creativity in self-organizing social networks. In *2021 9th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, pages 1–5.

[4] Bright, D., Sadewo, G. R. P., Lerner, J., Cubitt, T., Dowling, C., and Morgan, A. (2023). Investigating the dynamics of outlaw motorcycle gang co-offending networks: The utility of relational hyper event models. *Journal of Quantitative Criminology*.

[5] Byron, K., Keem, S., Darden, T., Shalley, C. E., and Zhou, J. (2022). Building blocks of idea generation and implementation in teams: A meta-analysis of team design and team creativity and innovation. *Personnel Psychology*, 76(1):249–278.

[6] Byron, K., Keem, S., Darden, T., Shalley, C. E., and Zhou, J. (2023). Building blocks of idea generation and implementation in teams: A meta-analysis of team design and team creativity and innovation. *Personnel Psychology*, 76:249–278.

[7] Chen, H., Mehra, A., Tasselli, S., and Borgatti, S. P. (2022). Network dynamics and organizations: A review and research agenda. *Journal of Management*, 48(6):1602–1660.

[8] contributors, W. (2024). Boardgamegeek – wikipedia, the free encyclopedia. `https://de.wikipedia.org/wiki/BoardGameGeek`. Accessed: 2024-07-27.

[9] Cortinovis, N. and van der Wouden, F. (2021). Better by design? collaboration and performance in the board-game industry. Papers in Evolutionary Economic Geography (PEEG) 2104, Utrecht University, Department of Human Geography and Spatial Planning, Group Economic Geography. Revised Jan 2021.

[10] Ertug, G., Brennecke, J., Kovács, B., and Zou, T. (2022). What does homophily do? a review of the consequences of homophily. *Academy of Management Annals*, 16(1):60–90. Published Online: 26 Jan 2022.

[11] Fleming, L., Mingo, S., and Chen, D. (2007). Collaborative brokerage, generative creativity, and creative success. *Administrative Science Quarterly*, 52(3):443–475.

[12] Harvey, S. and Berry, J. W. (2023). Toward a meta-theory of creativity forms: How novelty and usefulness shape creativity. *Academy of Management Review*, 48(3). Published Online: 31 Jul 2023.

[13] Hunicke, R., LeBlanc, M., and Zubek, R. (2004). MDA: A formal approach to game design and game research. In *Proceedings of the AAAI Workshop on Challenges in Game AI*, pages 1–5, San Jose, California. AAAI Press.

[14] Hâncean, M.-G., Perc, M., and Lerner, J. (2020). The coauthorship networks of the most productive european researchers. *Scientometrics*, 126(1):201–224.

[15] Ilgen, D., Hollenbeck, J., Johnson, M., and Jundt, D. (2005). Teams in organizations: from input-process-output models to imoi models. *Annual Review of Psychology*, 56:517–543.

[16] Juhász, S., Tóth, G., and Lengyel, B. (2020). Brokering the core and the periphery: Creative success and collaboration networks in the film industry. *PLoS ONE*, 15(2):e0229436.

[17] Klonek, F., Gerpott, F. H., Lehmann-Willenbrock, N., and Parker, S. K. (2019). Time to go wild: How to conceptualize and measure process dynamics in real teams with high-resolution. *Organizational Psychology Review*, 9(4):245–275.

[18] Lantz Friedrich, A. and Ulber, D. (2017). Why are we in a team? effects of teamwork and how to enhance team effectiveness. In Chmiel, N., Fraccaroli, F., and Sverke, M., editors, *An Introduction to Work and Organizational Psychology: An International Perspective*, pages 309–329. John Wiley & Sons, Ltd, third edition edition. First published: 11 March 2017.

[19] Lantz Friedrich, A., Ulber, D., and Friedrich, P. (2020). *The Problems with Teamwork, and How to Solve Them*. Routledge, 1st edition.

[20] Lerner, J. and Hâncean, M.-G. (2022). Micro-level network dynamics of scientific collaboration and impact: Relational hyperevent models for the analysis of coauthor networks. *Network Science*, 11(1):5–35.

[21] Lerner, J. and Lomi, A. (2023). Relational hyperevent models for polyadic interaction networks. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 186(3):577–600.

[22] Lerner, J., Lomi, A., Mowbray, J., Rollings, N., and Tranmer, M. (2021). Dynamic network analysis of contact diaries. *Social Networks*, 66:224–236.

[23] Levi, D. (2001). Introduction to group dynamics. In *Group Dynamics for Teams*, pages 1–24. Sage Publications, Inc, Thousand Oaks, CA, US.

[24] Linhardt, R. and Salas, E. (2023). Examining the fluidity of innovation teams: a conceptual framework. *Frontiers in Psychology*, 14:1296651.

[25] Liu, L., Jones, B. F., Uzzi, B., and Wang, D. (2023). Data, measurement and empirical methods in the science of science. *Nature Human Behaviour*, 7(7):1046–1058.

[26] Lutter, M. and Weidner, L. (2021). Newcomers, betweenness centrality, and creative success: A study of teams in the board game industry from 1951 to 2017. *Poetics*, 87:101535.

[27] Mannucci, P. V. and Perry-Smith, J. E. (2022). "who are you going to call?" network activation in creative idea generation and elaboration. *Academy of Management Journal*, 65(4):1192–1217.

[28] Paulus, P., Dzindolet, M., and Kohn, N. (2011). Collaborative creativity-group creativity and team innovation. *Handbook of organizational creativity*, pages 327–357.

[29] Perry-Smith, J. E. and Mannucci, P. V. (2017). From creativity to innovation: The social network drivers of the four phases of the idea journey. *Academy of Management Review*, 42(1):53–79. Published Online: 14 Oct 2015.

[30] Perry-Smith, J. E. and Shalley, C. E. (2003). The social side of creativity: A static and dynamic social network perspective. *Academy of Management Review*, 28(1):89–106.

[31] Pollok, P., Amft, A., Diener, K., Lüttgens, D., and Piller, F. T. (2021). Knowledge diversity and team creativity: How hobbyists beat professional designers in creating novel board games. *Research Policy*, 50(8).

[32] Rosing, K., Bledow, R., Frese, M., Baytalskaya, N., Johnson Lascano, J., and Farr, J. L. (2018). The temporal pattern of creativity and implementation in teams. *Journal of Occupational and Organizational Psychology*, 91(4):798–822.

[33] Shepherd, M. (2024). Recommend.games: Personalized board game recommendations. `https://recommend.games/`. Accessed: 2024-07-27.

[34] Soda, G., Mannucci, P. V., and Burt, R. S. (2021). Networks, creativity, and time: Staying creative through brokerage and network rejuvenation. *Academy of Management Journal*, 64(4). Published Online: 13 Sep 2021.

[35] Sousa, M., Zagalo, N., and Oliveira, A. P. (2021). Mechanics or mechanisms: defining differences in analog games to support game design. In *2021 IEEE Conference on Games (CoG)*, pages 1–8.

[36] Vedres, B. and Vasarhelyi, O. (2023). Inclusion unlocks the creative potential of gender diversity in teams. *Sci Rep*, 13(1):13757. Vedres, Balazs Vasarhelyi, Orsolya eng Research Support, Non-U.S. Gov't England 2023/08/24 Sci Rep. 2023 Aug 23;13(1):13757. doi: 10.1038/s41598-023-39922-9.

[37] Wachs, J. and Vedres, B. (2021). Does crowdfunding really foster innovation? evidence from the board game industry. *Technological Forecasting and Social Change*, 168.

[38] Woods, S. (2012). *Eurogames: The Design, Culture and Play of Modern European Board Games*. McFarland & Co, Jefferson, North Carolina. Accessed via EBL.

[39] Zhou, J. and Hoever, I. J. (2023). Understanding the dynamic interplay between actor and context for creativity: Progress and desirable directions. *Annual Review of Organizational Psychology and Organizational Behavior*, 10(1):109–135.