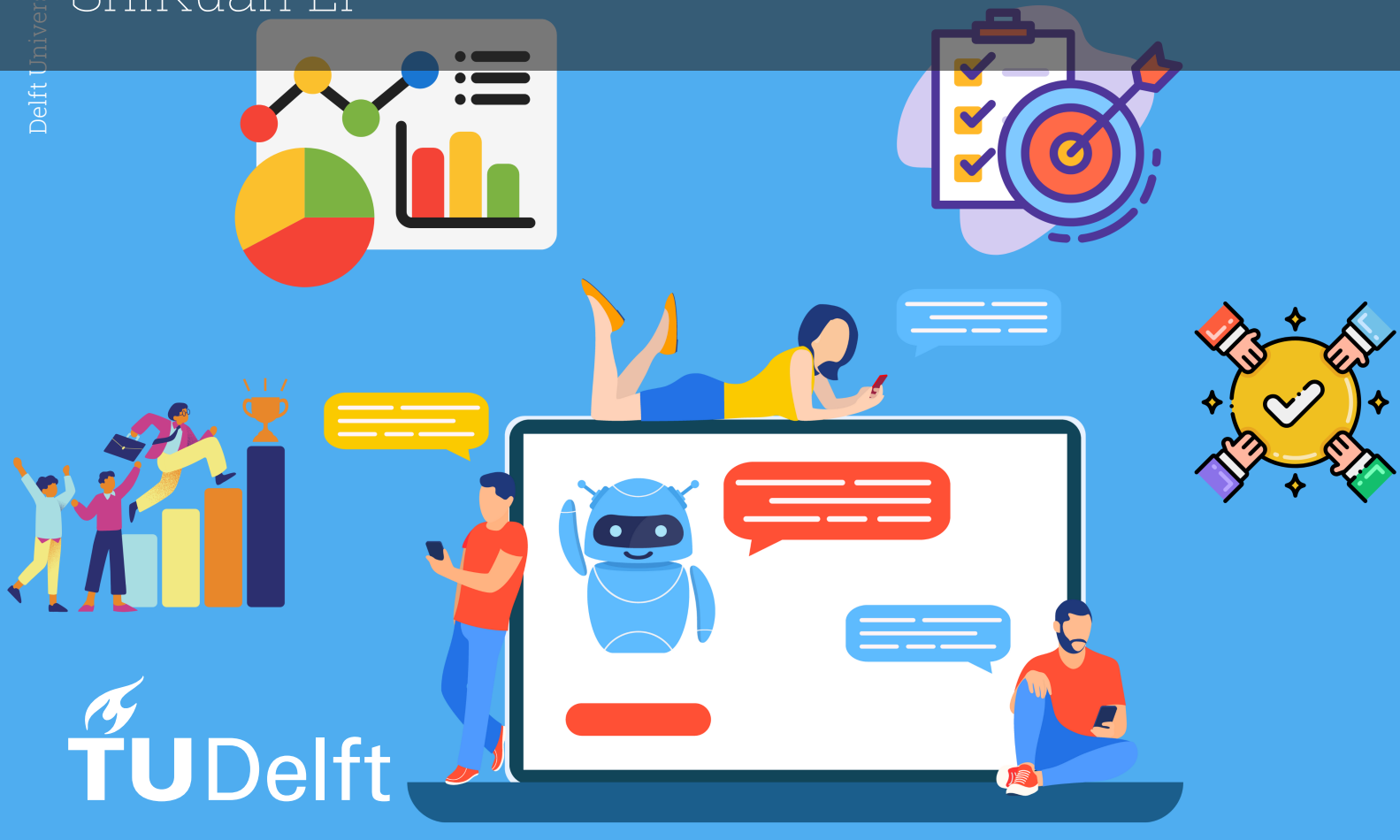
Optimizing Student Teamwork: Improving User Engagement and Collaborative Effectiveness

with SMART Collaborative-Goal-Setting Based Chatbot and Effort Visualization Tool

Master Thesis

ShiKuan Li

Delft University of Technology



TU Delft

Optimizing Student Teamwork: Improving User Engagement and Collaborative Effectiveness

with SMART Collaborative-Goal-Setting Based Chatbot and Effort Visualization Tool

by

ShiKuan Li

Student Name	Student Number
Shikuan Li	5621267

Instructor:	U. Gadiraju
Daily Supervisor:	T. Abbas
Project Duration:	November, 2022 - November, 2023
Faculty:	Faculty of Electrical Engineering, Mathematics and Computer Science, Delft

Style: TU Delft Report Style

TU Delft

Preface

This thesis represents a significant achievement during my pursuit of a Master's degree in Computer Science at the Delft University of Technology, in the Faculty of Electrical Engineering, Mathematics, and Computer Science (EEMCS). This research not only marks the pinnacle of my master's study journey but also embodies my commitment to academic rigor and dedication. Over the past year, I faced numerous challenges, yet I also experienced continuous learning and growth, culminating in a profound sense of accomplishment and satisfaction. I submit this thesis with deep gratitude, as it signifies an important milestone in my academic journey and stands as a testament to my ongoing passion for research.

I would like to express my special thanks to my supervisor, Professor Ujwal Gadiraju, for his invaluable support and thoughtful guidance throughout the research process. Professor Gadiraju not only granted me substantial academic freedom but also helped me navigate various challenges with his profound insights. His wisdom and meticulous guidance were crucial to the successful completion of this thesis.

Although this thesis marks a high point in my academic journey at Delft University of Technology, it is not the end. I look forward to applying the valuable knowledge and experiences gained at EEMCS in my future academic and professional endeavors. This application is not just an extension of my current learning but also a continuation of my lifelong journey of learning. I believe these experiences will lay a solid foundation for my future explorations and growth.

ShiKuan Li
Delft, November 2023

Abstract

With the growing importance of teamwork in higher education, effective communication and goal congruence have become vital in improving the effectiveness of student teamwork. This study aims to design and implement an innovative system that combines a goal-setting chatbot and an effort visualizer to facilitate effective collaboration in student teams. We found that unclear goals and lack of engagement were the main challenges in teamwork. To address these challenges, we developed a system that combines a chatbot and an effort visualizer designed to enhance team transparency and goal alignment through technology. The chatbot guides users to set specific collaborative goals through the SMART framework. At the same time, the effort visualizer displays each member's contribution, thus increasing accountability among team members and facilitating greater participation. The system we designed was validated in a creative writing task (N=84) in a collaborative scenario. The results showed that either the chatbot or the effort visualizer alone positively impacted student engagement and collaboration. However, when using the chatbot and effort visualizer in combination, there needed to be more clear evidence to improve engagement and collaboration effects further. This finding suggests that we pay more attention to the characteristics of individual tools and combination strategies when designing collaboration enhancement tools. This study provides valuable insights into using chatbots and visualization techniques to improve team collaboration and suggests new directions for future research in educational technology.

List of Figures

List of Tables

Contents

1 Introduction

1.1. Research Motivation

In the current rapidly changing educational environment, the importance of student teamwork has become particularly prominent. A study by Aranzabal et al. [4] delves into the critical role of collaboration in facilitating students' personal development and success in the modern educational environment. Effective teamwork enhances students' communication, leadership, and collaboration skills and is critical to their future career development. However, the current higher education system needs to improve in producing students with exceptional teamwork skills. Student teams often need help with communication barriers, misaligned goals, and social looseness during collaboration, which severely impediments to effective collaboration.

Student teams often face numerous difficulties and challenges during collaboration. Especially in teamwork, they are often in dire need of more initial experience and guidance. Differences in perspectives among team members sometimes lead to disagreements and cognitive dissonance, which not only impede the smooth progress of team tasks but also increase negative emotions and decrease overall work efficiency. As Martin noted in his 2022 study [29], these factors significantly affect the effectiveness of teamwork. In addition, inconsistency of goals among team members is another key factor affecting effective collaboration. Lin emphasized in his study that project failures often stem from a lack of clear goals to measure progress and poor communication, leading to differences in perceptions among team members. [25] Meanwhile, team transparency plays a crucial role in the collaboration process. Team tasks need to be more transparent for each member to understand what other teammates are working on, thus blurring the feeling of their contributions and inputs, creating barriers to collaboration. Therefore, there is an urgent need for a tool that helps student teams manage and align team goals and effectively demonstrates the extent of each team member's contribution.

Currently, widely used collaboration tools, including platforms like Trello, Miro, and Slack, play a significant role in assisting students with managing team projects[53]. These tools are primarily designed to monitor the accomplishments of each team member within set timeframes. However, there is scope for improvement in how they support the formation and relevance of team goals, which is essential for effective collaboration. One fundamental limitation is their focus on individual tasks rather than fostering a collective understanding of team objectives. Consequently, these tools may need to assess each member's contribution accurately or facilitate consensus on common goals. This shortfall can impede the overall effectiveness of team collaboration by not adequately addressing the dynamics of goal alignment and shared responsibility.

Additionally, several studies have pointed to chatbots as tools to facilitate in goal setting that can help teams clarify common goals and facilitate effective communication among members. Such tools are based on the SMART goal-setting theory, which helps teams set specific, measurable, achievable, relevant, and time-bound goals. This approach to goal setting helps team members clearly understand their tasks and responsibilities and dramatically reduces misunderstanding and confusion within the team. At the same time, effort visualization tools increase transparency, engagement, and accountability within the team by visually displaying each member's workload and contributions. This visualization tool motivates team members and promotes a sense of fairness and cooperation within the team. How-

ever, no research has yet explored the effects of using chatbots in conjunction with effort visualization tools to simultaneously achieve the dual functions of helping teams set collaborative goals and visualizing members' contributions.

Based on these considerations, this study proposes an innovative solution. We designed and implemented an integrated system incorporating a goal-setting chatbot and an effort visualization tool. The core goal of this system is to enhance intra-team communication by guiding teams to set collaborative goals and to drive team members towards goal congruence. In addition, through the integrated effort visualizer, the system can clearly show each team member's specific contributions, thereby increasing team transparency and engagement and reducing social looseness within the team. The novelty of this study is that it combines these two tools, aiming to investigate whether this fusion can further improve the efficiency and effectiveness of teamwork.

1.2. Research Question

The main goal of this thesis is to design and implement an innovative collaboration system designed to assist teams of students to explicitly set goals in collaborative tasks and effectively visualize each member's contribution level within the team. The core motivation of this research is to provide student teams with a more efficient mode of collaboration that motivates them to complete collaborative tasks more efficiently and achieve more profound engagement and collaboration effectiveness. Implementing such a system is expected to significantly optimize student teams' collaborative processes and enhance their overall performance. Thus, the main research question can be defined as follows:

Research Question 1 **How can we design an interactive system to promote student collaboration?**

To answer this question, we conducted an in-depth literature review focusing on the practice and effectiveness of various tools currently used to intervene in team collaboration and shed light on the key factors affecting collaboration effectiveness. In the literature review section, we summarize in detail the process of using chatbots to guide goal setting, and analyze in detail the specific needs for designing such chatbots based on goal setting theory and collaborative goal setting. Similarly, we synthesize the existing literature, provide an overview of the role of visualization tools and their contributions to collaborative tasks, and explore how to design an effective visualization tool designed to demonstrate the contributions of team members.

To comprehensively evaluate our collaborative system, we employed an empirical research methodology for testing. Specifically, we selected a creative collaboration task as the backdrop for assessment, aiming to explore team performance and effectiveness when utilizing this system. Upon completion of the collaborative task, we gathered data on user engagement and collaborative effectiveness from the participants, thereby observing the system's impact on team collaboration.

This assessment process also aimed to address our second research question

Research Question 2 **How do different features of the collaborative system influence user engagement and collaboration effectiveness?**

We focused particularly on how various aspects of the system affect team members' interaction styles and the effectiveness of their collaboration. Through this empirical research approach, not only can we assess the current effectiveness of the system, but also explore potential areas for optimization. Such analysis will directly guide our future refinements of the system, aiming to maximize its positive impact on team collaboration. The overarching goal of this research is to explore and realize more efficient and creative modes of team collaboration, thereby advancing the state of team cooperation at a technological level.

1.3. Approach

To address the research questions above, the first and most crucial step in our study was to conduct a comprehensive literature review, aiming to understand the current research and available solutions in this field. In Chapter 2, we extensively summarized the importance of collaboration and the key factors influencing student collaboration through a review of the literature. At the same time, we explored the advantages brought by the use of visualization technologies in collaborative processes. Subsequently,

we summarized the use of chatbots related to goal setting and collaboration, delving into goal-setting theory and collaborative goal setting. This chapter aimed to understand how to design a system for enhancing collaboration efficiency.

Following this, in Chapter 3, based on the findings from our literature review, we established a collaborative system in a creative writing context, named TeamSynerPad. We detailed how we designed and implemented this system aimed at improving collaboration. In Chapter 4, we conducted a between-group design experiment to assess the performance of this system in enhancing collaboration and discussed the results in Chapter 5. Chapter 6 delved deeper into the system's limitations and considered participants' feedback and future research directions. Finally, in Chapter 7, we concluded the study by summarizing the findings and presenting our conclusions.

2

Related Work

2.1. Collaboration among students

Student collaboration plays a crucial role in higher education, and numerous studies have explored this in depth from various perspectives. Pilar's study highlights the importance of teamwork skills in developing innovative and adept teamworkers [37], particularly the positive impact of teamwork on enhancing the effectiveness of collaboration. Similarly, Andreea's study indicates the importance of teamwork for collaboration[14], especially the need to develop communication skills in learning through collaboration. In addition, researchers such as Jade and Eva conducted an in-depth study on students' perceptions of collaborative group learning (CGW) [30] [16]. They found that collaborative group learning increased student engagement and developed a variety of soft skills such as communication, teamwork, and conflict resolution. These studies show that effective teamwork communication is critical to successful student engagement. Students' active participation and efficient communication in teamwork significantly improved learning outcomes and overall team performance. On the other hand, Alexandra's study highlights the importance of online collaboration skills in the current era of globalization and digitization. Further, it emphasizes the centrality of communication and participation in teamwork in cross-cultural and digital environments[21]. Together, these studies suggest optimizing communication and increasing student engagement are key factors for effective collaboration in higher education.

However, existing research also reveals challenges in student collaboration. Aranzabal identifies challenges in student teams' ability to communicate and work effectively[4], especially when they need more experience and support for teamwork. Hung and Vitaliy explored challenges, such as communication difficulties and low personal responsibility, which negatively impact collaboration [49] [36]. Helen et al. focused on the challenges of online team projects and teamwork as a learning incentive, emphasizing the negative impact of a lack of clarity of purpose and consistency in decision-making processes among team members on team performance and cohesion [10] [16]. Together, these studies reveal complex challenges in student collaboration that must be overcome through effective communication, collaborative goal-setting, and team-building strategies.

Factors influencing student collaboration and strategies for achieving effective collaboration include several key elements. First, the active participation of team members and practical construction between teams are critical to collaboration success [37]. Hoi et al. emphasized the importance of predicting team performance in collaborative project-based learning, and they highlighted the impact of clear goal-setting and effective communication on team success, as well as the need for teams to be clear about their goals and to maintain transparency in communication to increase efficiency [25]. In addition, Hung's study explored member perceptions from the perspective of collaborative online team learning, highlighting the importance of increasing transparency and team cohesion [49]. Maintaining strong connections and clear communication channels among team members is critical to collaboration success in digital learning environments [49]. Further, several studies highlighted the importance of several other key elements for improving the effectiveness of student collaboration. Yao emphasized the value of using learning analytics tools, such as Slack, to support collaborative learning [53], demonstrating the important role of modern technological tools in facilitating effective communication and collaboration among students. Shunan explored the effect of familiarity among team members on in-

creasing collaboration satisfaction, pointing out that relationships among team members can influence team effectiveness zhang2023know.

Furthermore, Amuanuel et al. emphasized the importance of effective conflict management for team success by examining team conflict, conflict management, cohesion, and effectiveness [45]. This finding highlights the importance of how internal disagreements and conflicts are handled in teamwork. Finally, Adesina explored the impact of peer assessment on student engagement, emphasizing the role of mutual evaluation and feedback within a team in enhancing the effectiveness of collaboration [2]. Taking these studies together, we can see that team engagement, clarity of goal setting, effective communication, and team cohesion are key factors in the success of student collaboration.

2.2. Visualization Techniques in Collaboration

Visualization technology is critical in facilitating teamwork, as demonstrated in several studies. Ella Glikson's effort visualizer highlights the value of visualization systems in providing real-time feedback on team members' behaviors and that this immediate feedback can be used as an intervention tool to reduce the loss of process in a team's efforts, thereby increasing team member self-awareness and overall team Collaborative Efficiency[15]. Fernanda's research combines computer-supported collaborative work (CSCW) with information visualization, exploring how these technologies can be used to support collaborative decision-making, especially in situations where different work schedules may lead to conflict[32].

Rupayan Neogy's work addresses a key challenge in collaborative visualization: building common ground and maintaining a sense of sharing among participants. His proposed model enhances engagement through an interactive approach that allows users to quickly view the work of collaborators through different perspectives, facilitating more profound and more efficient collaboration [32] Meanwhile, the digital dashboards developed by Rohan et al. provide real-time monitoring and evaluation of student teams, and this monitoring tool not only facilitates the completion of tasks by team members but also increases their commitment to the team's process and the project [3].

Isenberg's study defines the new research field of collaborative visualization. It highlights how much easier it has become for multiple people to share and collaborate through the Internet, mobile devices, and shared displays. These advances make it easier for multiple people to share and analyze information and form decisions together [19] Visualization techniques for learning analytics, such as those explored in Sanam's study, have also demonstrated their potential to support the learning process, especially when it comes to understanding students' achievement goal orientations [6].

On the other hand, Xi et al.'s study focuses on Slack-enabled collaborative learning, revealing the importance of visualization tools in measuring and enhancing student engagement [53]. Engagement corresponds to the role of familiarity among team members, and engagement is a critical factor in enhancing the satisfaction and efficacy of teamwork [52]. In addition, teams performed better than individuals on tasks requiring pooled thinking, highlighting the importance of shared goals in promoting creativity and optimizing organizational work team decision-making [46].

Overall, these studies suggest that visualization techniques enhance the transparency and efficiency of teamwork and increase familiarity between teams by increasing team member engagement and satisfaction, ultimately enhancing overall team performance.

2.3. Existing chatbot for Education And Collaboration

The integration of chatbots and education has changed traditional teaching methods and ushered in a new era of personalized and interactive education. Chatbots, also known as conversational agents, enable humans and computers to interact through natural language by applying Natural Language Processing (NLP) techniques [22]. The study by Wollny et al. emphasizes the role of chatbots in "teaching," "assisting," and "guiding," showing their usefulness in improving collaboration problems among students. Wollny et al. emphasized the role of chatbots in "teaching", "assisting," and "guiding," showing their potential to improve collaboration problems among students [50]. IVO BENKE et al. present chatbots for managing teamwork, showing that teams that communicate with chatbots demonstrate greater communication efficiency and cohesion [7].

IntroBot was designed to utilize social media data to help team members become familiar with each other, increasing trust and cohesion within the team [40]. In educational settings, the use of chatbots has been found to enhance students' goal-setting skills and increase engagement and motivation

[17][13]. Weijiao's study showed that chatbots acted as goal-setting facilitators in online courses, improving the learning experience for students [18]. In addition, rule-based systems for chatbots are widely used to improve user engagement and provide personalized experiences [47].

In the context of teamwork, Susa, a conversational agent developed by Sofied et al. was used to facilitate teamwork and collaborative practices [42], while a chatbot designed by Mateus et al. assisted in coordinating group tasks [41]. In practical collaboration, the chatbot proposed by Amon et al. emphasizes practical collaborative strategies executed in human-computer interaction [39].

Chatbots improve student engagement and learning outcomes in flipped learning environments by providing personalized support, facilitating group discussion and collaboration, and providing feedback and assessment of student work [5]. Chatbots reduce isolation and enhance a sense of ”belonging” for mature students' online engagement [1]. On the other hand, Differ, a communication platform with chatbot functionality, was used to increase student engagement and course completion [9].

A study by Jeya et al. investigated the impact of integrating an educational chatbot (EC) to facilitate a team project for designing a course on learning outcomes and found that EC improves academic performance and teamwork [23]. Finally, the introduction of chatbots in educational environments requires careful and thoughtful strategies, such as the role of ChatGPT in providing instant feedback, query response, and real-time help, which offers the potential to overcome language barriers, assist students with disabilities, and facilitate collaborative learning experiences [38].

In summary, chatbots are increasingly being used in educational and collaborative environments, and their contribution to enhancing the student learning experience, increasing team cohesion and efficiency, and facilitating personalized and interactive education cannot be ignored. By incorporating natural language processing technologies, these tools not only increase the accessibility and personalization of learning but also greatly enhance student learning outcomes and team collaboration by supporting goal setting and teamwork. As the technology develops further, we can foresee chatbots playing an even more critical role in the future of education and teamwork.

2.4. Goal-setting Theory

From the above literature analysis, using chatbots in collaborative environments can be effective in improving communication and helping teams make progress in goal-setting. In order to gain a deeper understanding of the role of chatbots in the goal-setting process, it becomes crucial to explore goal-setting theory. It would provide us with a theoretical basis and guidelines for designing a chatbot specializing in assisting users to set goals.

Locke and Latham are leaders in the study of goal setting and motivation theory [27]. Their work dates back to the 1960s, and this collaborative research laid the foundation for goal-setting theory, which has continued to be developed and refined over the ensuing decades. According to Locke and Latham, in order to motivate work and improve performance, goals should have specific characteristics, including clarity, specificity, challenge, and measurability [26]. These characteristics help individuals understand goals, measure progress, and feel challenged, which enhances work motivation. Goal-setting theory emphasizes that when individuals set clear goals, and these goals are aligned with their values and expectations, they are more likely to engage in positive behaviors to achieve these goals [28]. Such positive behaviors include more productive work and more focused attention. At the same time, goal-setting theory emphasizes the importance of feedback. Individuals adjust their behavior by obtaining information about progress toward goal attainment. Positive feedback can enhance work motivation, while negative feedback may lead individuals to readjust their goals or strategies [26]. In addition, moderately challenging goals may inspire higher performance, but overly complex goals may lead to frustration and discouragement[27]. Goal setting applies not only to individuals but also to organizations. By establishing clear organizational goals and communicating them to members within the organization, the performance and effectiveness of the entire organization can be improved.

Based on the goal-setting theory, the SMART (Specific, Measurable, Achievable, Relevant, Time-bound) goal-setting framework has been introduced and widely embraced across various domains [11] [8]. This framework underscores the critical attributes of goals to ensure their effectiveness. The following elucidates each constituent of the SMART framework in detail:

- **Specific**: The goal needs to be clearly defined, describing in detail the content, scope, action plan, and desired outcome of the task. A clear goal helps people to understand and act upon it more clearly.

- **Measurable:** Goals should be measurable to evaluate progress and results. It requires well-defined indicators and metrics that help to monitor the extent to which the goal has been achieved.
- **Achievable:** Goals should be realistic and feasible, considering the goal setters' resources, skills, and conditions. Goals should be challenging but also avoid being too difficult to maintain motivation.
- **Relevant:** This component refers to ensuring the relevance of the goal to the specific task or project to ensure that resources and efforts are utilized efficiently and to improve the chances of success.
- **Timebound:** Goals need clear deadlines to create a sense of urgency, prevent procrastination, and encourage timely action.

2.5. Collaborative goal-setting

After establishing the fundamental principles of goal-setting theory, our attention turns towards its practical application within group settings. Given the specific focus of our study on goal setting within collaborative teamwork, this section delves into the concept of collaborative goal setting and provides a summary of the factors influencing goal setting in teams. This analysis is grounded in Locke and Latham's goal-setting theory [28].

Collaborative goal setting is a way of identifying and establishing group goals. Typically, the process involves identifying individual and group-level requirements and evaluating goal performance over time. According to research, collaborative team goals enhance interdependence and collaboration, facilitating teamwork and achieving team goals [24]. In addition, participants who set collective goals successfully exceeded their individual goals [26]. The target audience for collaborative goal setting is students working collaboratively on group projects in any college-level course [37]. In a group project, members set a variety of goals and then work to achieve those goals to maximize individual and group outcomes [18]. To ensure fairness in the distribution of effort [31], we expect each member of the process to make an equal contribution][31], which is difficult to evaluate objectively in practice. In addition, members are sometimes unaware of their key roles and responsibilities, which is what prevents the group from working together effectively and why we use collaborative goal setting to remove this impediment.

At the same time, Locke and Latham point out that individuals do not operate independently in a team-based structure. Instead, they must engage in interpersonal interactions and other processes to accomplish the team's goals. For example, completing a team task requires various processes, such as communication and coordination, that do not exist in individual tasks. A team is defined as a group of two or more people who interact dynamically, interdependently, and adaptively, working toward a common, valued goal/task/mission, assigned a specific role or function to fulfill [28]. At the same time, teams are formed to capitalize on their members' unique knowledge, skills, and abilities. Without clear goals (as defined by the team's objectives), there are no markers to guide the definition of roles or to determine how to participate in the team process, which is one of the causes of team failure. Research has shown structural differences in goal setting between individuals and teams [51]. In a team, each member may have goals for his or her performance and larger goals for the team. Hence, ensuring that each member's individual goals are aligned with the team's collaborative goals is a prerequisite for good collaboration. Achieving consistency between individual and team goals is also a process of achieving team awareness, cohesion, and trust. Good team alignment can motivate teams to work towards collaborative goals [43]. James' research suggests a process for setting collaborative goals that includes clarifying task requirements, building team consensus and setting team goals, and coordinating and monitoring the tasks assigned to each team member [12]. The study also found that setting collaborative goals moderated the quality of teamwork and that reasonable team goals and a clear division of labor could lead to a higher quality of collaboration.

3

Study Design

In this chapter, we delve into the detailed design and implementation of the system. Initially, Section 3.1 articulates our rationale behind choosing creative writing as a collaborative task. A comprehensive overview of the main system is shown in Section 3.2, including the competent design and implementation. After that, Section 3.3 elucidates the dialogue design of the collaborative goal-setting chatbot.

3.1. Creative Writing Task

Collaborative tasks among students are varied because different programs and areas of specialization have different requirements, and students need to respond to various tasks. These tasks range widely, including writing code together, completing assignments together, and solving problems. However, we chose the creative writing task as a typical example because students collaborating on reports is a common team task in many courses and fields. This task requires students to think together, create collaboratively, integrate ideas, and ultimately present a comprehensive document that is important for developing teamwork, creative thinking, and communication skills. In the meanwhile, Choosing creative writing tasks as collaborative task offers several advantages:

- Interdependence: Collaborative creative writing tasks require participants to work together to achieve a common goal. Each group member depends on the others for input, feedback, and support. By working together, they can create a better story than they could have created individually.
- Communication: Collaborative creative writing tasks require effective communication skills. Participants must be able to articulate their ideas and listen to the ideas of others. Working together to negotiate and refine ideas can create a more cohesive and effective outcome.
- Conflict resolution: Collaborative creative writing tasks can also test a group's ability to resolve conflicts. Disagreements and differences of opinion are inevitable in a group setting, but the ability to work through these conflicts and find a resolution is critical for success.
- Teamwork: Creative writing tasks require teamwork, which involves cooperation, mutual support, and a shared commitment to the group's goals. Participants can develop important leadership, accountability, and adaptability skills by working together.

3.1.1. Creative Writing Scenarios

Topic 1: Time Traveler

Welcome to a fantastic journey across time and space! In this scenario, students will delve into the mysterious journey of time, encountering the legacies of the past or the wonders of the future. Time travelers' stories involve elements of science or fantasy and deep philosophical and moral reflections.

Your task is to write a story about time travel in a time machine. Please consider the following points:

- What is the identity and provenance of the time travelers?
- What is their purpose in traveling through time?
- What are the challenges and opportunities they will face in time?

- How is the continuity of time and possible paradoxes skillfully handled?
- Does the story incorporate historical or futuristic elements to enrich the characters' adventures?

Please make sure you weave a creative story that piques the reader's interest and explores the possible effects of time travel on individuals and society.

Topic 2: Alien Civilization

Space is a vast and infinite realm of unknown and exotic possibilities! In this topic, you will examine the existence of alien civilizations, life forms that may be vastly different, with unique cultural, technological, and philosophical concepts.

Your mission is to construct a story involving human interaction with an alien civilization. Consider the following aspects of this cosmic adventure:

- What are the characteristics of the alien civilization? What is their technology, culture, and social structure?
- What was the first contact between humans and the alien civilization like? Was it a friendly exchange or an outbreak of war?
- What are both civilizations' misconceptions and preconceptions, and how do these factors affect their interactions?
- Are there specific scientific, philosophical, or moral challenges and explorations in the story?
- Are there thoughts about human nature or the nature of extraterrestrial beings involved?

Please incorporate reflections on science, philosophy, and sociology into the story and make alien civilizations an exciting way to explore this topic.

We chose these two topics because the scenarios cover various fields, including science, history, philosophy, and sociology, so students can work together to explore and solve problems through an interdisciplinary approach. Meanwhile, it is an excellent opportunity for students to practice analyzing problems from different perspectives, integrating multidisciplinary knowledge, and working together in report writing.

3.2. Design overview and implementation detail:

In support of the collaborative task we are studying, creative writing, we have developed a collaboration system called TeamSynerPad. The system is based on the Etherpad editor, a Node.js web application. A1 and integrates an Effort Visualizer A2 and a Chatbot A3 for setting collaborative goals.TeamSynerPad aims to optimize collaboration and communication between teams by providing a real-time online writing platform (powered by Etherpad) while injecting innovative features to enhance the efficiency and synergy of teamwork.

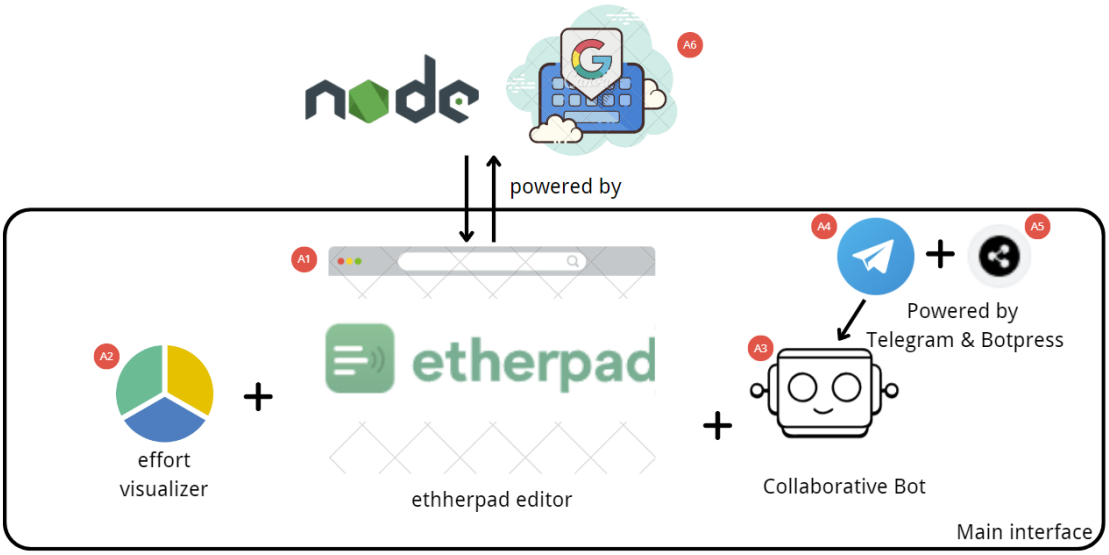


Figure 3.1: TeamSynerPad System Overview

3.2.1. Etherpad Editior

We considered the following key elements when choosing Etherpad Editor A1 as the basis for TeamSynerPad. Etherpad Editor [44] offers powerful real-time collaboration features that allow multiple users to work on a single document simultaneously. This real-time interactivity is a core requirement for creative writing team collaboration, and it enhances communication and collaboration efficiency between teams by displaying each team member's edits in real-time. Meanwhile, as an open-source project, Etherpad Editor is highly customizable. We can customize and extend the editor according to the specific needs of the project, such as integrating Effort Visualizer and chatbot, to achieve the functions that better meet our project needs. In addition, Etherpad Editor offers simple deployment options and can be easily integrated into our Google Cloud platform A6. This ease of deployment and integration lowered the technical barriers to system development and maintenance, allowing us to launch and test new features and improvements quickly. Finally, Etherpad Editor has excellent data persistence and version control to record the history of each edit, which is important for creative writing teams because it allows team members to go back and compare different versions of a document better to understand the context and background of each edit.

3.2.2. Effort Visualizer

In TeamSynerPad, the 'Effort Visualizer' plugin, highlighted as A2, significantly enhances team collaboration by providing a real-time visualization of each member's contribution in shared documents. Developed in JavaScript and seamlessly integrated into the Etherpad Editor, this innovative plugin operates by monitoring document edit events, which allows it to accurately capture and count the number of words each team member inputs. Contributions are then calculated as a percentage of the total word count and represented in a dynamic pie chart. This intuitive visualization updates in real-time as team members add or edit content, offering a clear and continuous measure of individual engagement. Such a feature not only fosters transparency but also aids in maintaining collaborative efficiency by keeping

the team updated on the latest contribution levels.

The design and functionality of the 'Effort Visualizer' are deeply influenced by the principles from Ella Glikson's research on *Visualized Automatic Feedback in Virtual Teams* [15]. Glikson's study underscores the importance of quantifiable and visual feedback in enhancing the collaboration and performance of virtual teams. Embracing these insights, the Effort Visualizer in TeamSynerPad uses a straightforward yet effective metric—the percentage of words contributed by each team member—to visualize contributions. By quantifying each member's input in this manner, the Effort Visualizer not only tracks individual contributions accurately but also encourages balanced participation.

3.2.3. Collaborative Chatbot

Subsequently, in TeamSynerPad, our chatbot A3 was implemented based on Botpress A5 and Telegram A4. We chose Botpress [48] as the development framework for our chatbot because it is an open-source, customizable, and powerful platform for building chatbots. Botpress provides a rich set of tools and interfaces, including a built-in Natural Language Processing (NLP) module, which helps us to understand and parse user inputs, thus enabling the chatbot to provide more accurate and relevant responses. In addition, the flexibility and customizability of Botpress allow us to easily define and configure the rules engine for the chatbot to implement our desired interaction logic, the specific session logic of which will be further discussed in section 3.3. In order to provide a real-time, easy-to-use user interaction interface, we chose to integrate the chatbot into the Telegram messaging platform. Telegram is known for its real-time messaging mechanism and efficient APIs, providing users with a familiar and user-friendly interaction environment [20]. Using Botpress' Telegram Integration Module, we created a chatbot on Telegram and provided TeamSynerPad with an icon to enter a Telegram group. Users can click this icon to go directly to a specific Telegram group and interact with the chatbot. This integration provides an intuitive user interface and makes it easy for users to interact with the chatbot while editing documents, thus increasing the efficiency and convenience of collaboration. Our chatbot is rule-based, where the rule engine is configured through the built-in functionality of Botpress. This rule engine can determine the chatbot's response based on predefined rules and user input, thus realizing our desired interaction logic. This rule-based design allows the chatbot to respond to user requests predictably and consistently.

With this design and implementation, TeamSynerPad aims to improve collaboration effectiveness, optimize work distribution, and facilitate team communication.

3.2.4. Interface

After participants decide to join the study, they will work in groups of three to complete the creative writing task. First, they will receive a link for accessing TeamSynerPad. Figure 3.2 illustrates the main entry screen of TeamSynerPad, where participants' groups of three will be asked to enter a Pad name that is specifically assigned and unique to them to access their creative writing space. This design ensured that each group could collaborate in a separate and dedicated online space, free from external interruptions, as well as ensuring the privacy and security of the content of their work.

When participants enter the creative writing space, they are presented with a detailed view of the inside of the TeamSynerPad platform, shown in Figure 3.3. In the initial section of the page, participants will find essential explanations and guidelines for the creative writing assignment, including an overview of the two optional topics, "Time Travelers" and "Alien Civilizations" to guide them as they begin to think about and shape their stories. Next, in a designated area, teams will work together to complete their collaborative goals, identifying the various milestones and expectations they want to meet as they move through the creative process. Once the collaborative goals are finalized, the team will move to the dedicated writing area and begin their creative journey to bring their ideas and stories to life on virtual paper.

Additionally, the Effort Visualizer displays each team member's participation and input in a pie chart, providing an intuitive way to track and reflect on the contributions and activities of each team member. Finally, Chatbot exists in the lower right corner of the interface, and members interact with Chatbot by clicking on the Chatbot icon to access the telegram group, designed to guide the team through a collaborative goal-setting process that provides ongoing direction and support for their writing tasks. These components build a versatile collaborative space to support and enhance the team's working experience on creative writing tasks.

Figure 3.2: Entrance page of Etherpad

3.3. Chatbot Dialogue Flow Design

	Requirements
Specific	The chatbot should provide the user with a choice of topics to select from in the creative writing.
Measurable	The chatbot should provide users with measurable indicators in creative writing, such as word counts and chapter counts.
Achievable	The chatbot should be able to make the user aware of whether or not they can achieve this goal.
Relevant	The chatbot should be able to provide users with examples to guarantee that their writing content aligns with the topic they choose.
Time-bounded	The chatbot should be able to give user feedback regarding the timing of the task.

Table 3.1: SMART requirements for collaborative-bot

To address RQ1, the system should include a chatbot to help student teams work together to set collaborative goals during creative writing tasks. This section will describe how we designed a conversational flow based on the SMART framework to guide student teams in setting collaborative goals. In addition, we will also explain how this chatbot facilitates team communication and ensures goal alignment in collaborative scenarios.

In this study, we emphasize the importance of collaborative goal setting as it is one of the critical factors in facilitating efficient group collaboration. At the same time, appropriate goal setting can help students improve team participation and trust. As mentioned in Chapter 2, a series of studies have shown that using the SMART framework can set collaborative goals. Therefore, we designed a conversational flow based on the SMART framework to guide the goal-setting process. In this process, we ask questions to users based on the different dimensions of the SMART framework (Specificity,

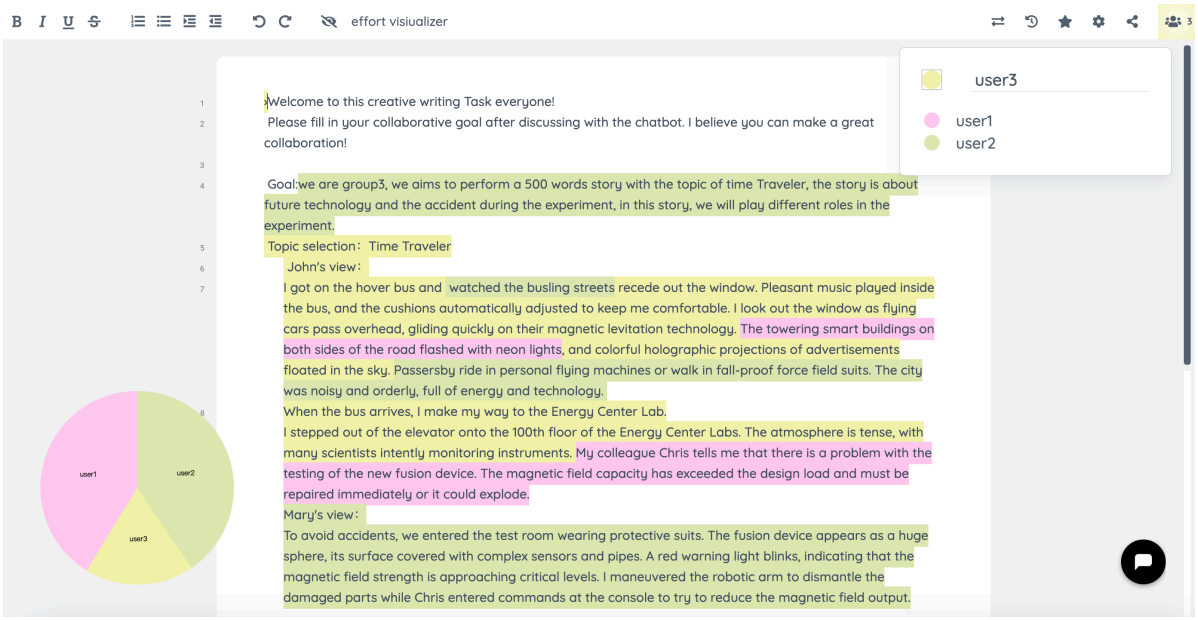


Figure 3.3: Main interface of TeramSynerPad

Measurability, Achievability, Relevance, and Timeliness) and ultimately integrate the results to align with the goals of the SMART framework.

Table 3.1 lists the requirements corresponding to each SMART goal in a creative writing scenario. Based on this framework, we implemented the conversation flow of our chatbot, illustrated in Figure 3.4.

First, in Step 1, the chatbot greets the user and introduces its function and the purpose of the conversation to help the team clarify the creative writing goal. In Step 2, the chatbot provides the user with examples of topics to choose from in creative writing and asks them to select to ensure their writing goals are specific and clear. Next, in Step 3, we implemented a polling feature in the chatbot that allowed the group to vote to select topics for writing. The reason for making a vote is that in collaborative scenarios, each team member may have different ideas, and voting is an effective way to build team consistency and help the team reach consensus quickly [40]. Once the team has finalized the writing topic through voting, we must ensure their goals are measurable. Therefore, in Step 4, the chatbot shows an example of how to guide the team to specify the number of words and chapters required for creative writing.

Subsequently, to guide teams in checking whether their goals are achievable, we designed Step 5, the Confidence Feedback Scale, which contains seven levels ranging from 1 (very unconfident) to 7 (very confident). We designed this feature to ensure team consistency in a collaborative environment, as different team members may need to be more consistent in assessing the difficulty of accomplishing the current goal. Team members can assess whether the goals set by the team are reasonable by observing how confident each individual is about completing the current task and adjusting the collaborative goals based on the feedback. After the team has clarified the achievable goals, the chatbot will guide them to ensure they are relevant. In step 6, the chatbot demonstrates the help provided based on the team's current topic selection to ensure that their writing content is relevant to the selected topic. At the same time, the chatbot encourages them to rationalize the division of tasks to ensure that each team member is involved in the collaborative task.

Finally, to ensure that collaborative goals were time-bounded, we request participants involved in creative writing tasks to form the overall collaborative goals of their group based on their interactions with the chatbot and document them in the etherpad editor. In addition, the chatbot can prompt the team every 20 minutes as checkpoints for task completion. Together, these steps ensure that collaborative goals are Specific, Measurable, Achievable, Relevant, and Time-Bounded, thereby enhancing the collaborative effectiveness of student teams in creative writing tasks.

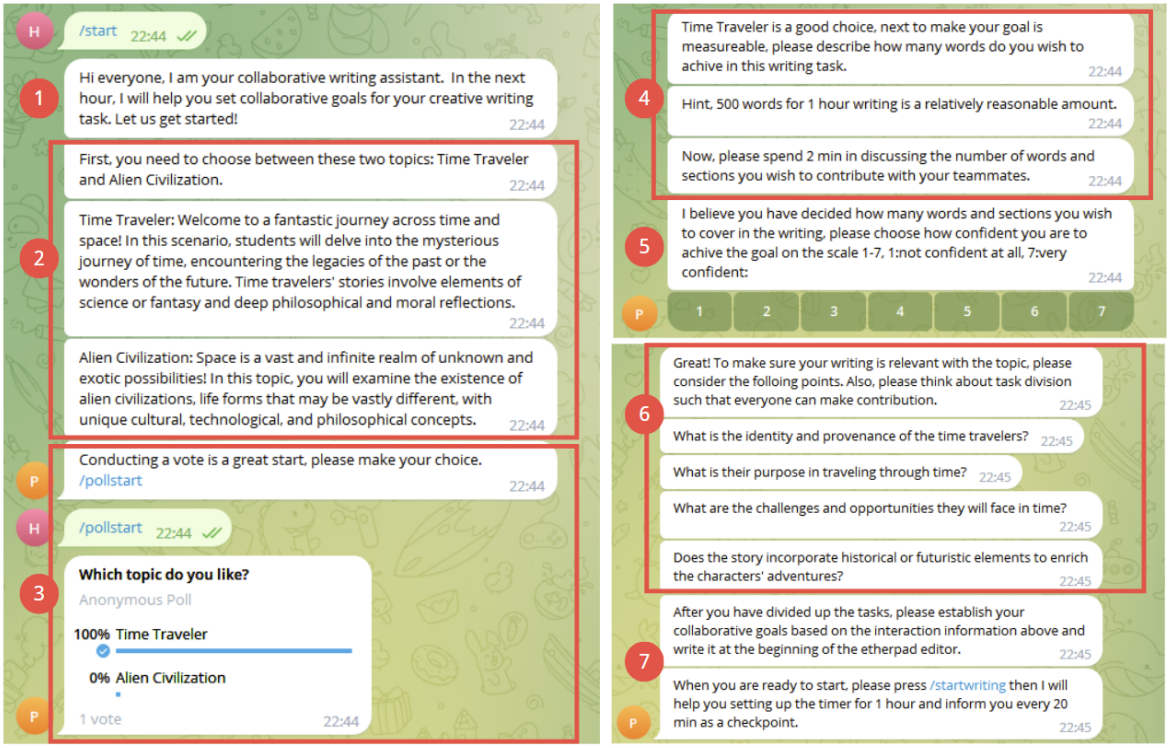


Figure 3.4: Chatbot Dialogue flow

4

Evaluating TeamSynerPad System

In this chapter, we first outline the experiment's goal and hypothesis in Section 4.1. Following this, Section 4.2 delves into our experiment design in detail, elaborating on a between-group design with four conditions that utilize different functionalities of TeamSynerPad. Sections 4.3 and 4.4 detail the participant requirements and the procedure for participating in the study, respectively. Section 4.5 discusses how user engagement and collaboration effectiveness were obtained through questionnaires. Finally, Section 4.6 describes the data processing approach, focusing on how we handled the collected data.

4.1. Hypotheses

In the previous sections, we explained the design of the TeamSynerPad System for promoting student collaboration in a creative writing scenario. As previously explained, no research combines different tools to promote student teamwork to explore whether better engagement and collaborative effectiveness can be achieved. Therefore, to address research question 2: How does using different features of TeamSynerPad make a difference in the overall collaboration effectiveness with regards to:

- user engagement
- collaboration quality
- quality of creative writing

Reflecting on the content of Chapter 2, the studies conducted by Shin[40], Hew[17], and Abbasi et al. illustrate how the use of chatbots in collaborative scenarios can enhance communication and student collaboration. These studies emphasize the role of chatbots in setting collaborative goals and increasing student engagement in higher education. Similarly, the research by Glikson [15] and Beheshitha et al. [6] demonstrates the benefits of using visualized dashboards to increase team transparency and engagement, thereby improving collaboration quality. Motivated by these findings, we also aim to investigate the combined impact of using both chatbots and visualization tools on student engagement and collaboration effectiveness. Based on these insights, we propose the following hypotheses:

(H1a) Using a chatbot positively impacts user engagement.
(H1b) Using the effort visualizer positively impacts user engagement.
(H1c) The use of both chatbot and effort visualizer has a more significant impact on user engagement than using only one tool.

(H2a) Using a chatbot positively impacts collaboration effectiveness.
(H2b) Using an effort visualizer positively impacts user engagement.
(H2c) The use of both chatbot and effort visualizer has a more significant impact on collaboration effectiveness than using only one tool.

(H3a) Using a chatbot positively impacts Creative Writing Quality.
(H3b) Using the effort visualizer positively impacts Creative Writing Quality.
(H3c) The use of both chatbot and effort visualizer has a more significant impact on user engagement than using only one tool.

4.2. Method

Our experiments aim to test the above hypotheses by exploring the impact of combined tools in promoting student collaboration. We assumed that combining these two tools would lead to a richer interactive experience for students in creative writing, which would influence their level of collaboration and lead to collaborative tasks in an organized way with this new system. We expected that using TeamSynerPad would result in more effective student collaboration.

Experimental Design A control experiment was adopted using a between subject design with 4 conditions. Figure 4.1 provides the overview of the 4 conditions that use different features in TeamSynerPad system.

Figure 4.1: Experimental Conditions

- Condition1: Under this condition, participants could use all the features of the TeamSynerPad system. It means that they can take use of the collaborative editor(Etherpad-light), effort visualization tools, and chatbot provided by the system to assist them in their creative writing tasks.
- Condition2: In this condition, participants could only use the effort visualization tool in the TeamSynerPad system. It means that they could see the contributions of their team members, but could not use the chatbot function. To help them set collaborative goals, we provided a basic guideline based on the SMART framework.
- Condition3: In this condition, participants could only use the chatbot in the TeamSynerPad system. The chatbot could help them set collaborative goals and provide checkpoints, but they could not use the effort visualization tool.
- Condition4: In this condition, participants could only use the basic editor functionality of the TeamSynerPad system, not the effort visualization tool and chatbot. To help them set collaborative goals, we provided a basic guideline based on the SMART framework.

To ensure the experiment's fairness and the results' reliability, we used randomization to assign participants to one of the four conditions. In determining the sample size, we considered the statistical efficacy of the experiment and available resources. Using G*Power's efficacy analysis, we determined the number of participants needed for each condition to ensure that the experiment had sufficient statistical efficacy to detect the intended effects. The experimental process began with an introductory and training phase for participants to ensure that all participants understood the TeamSynerPad system and the experimental task. Subsequently, participants in each condition were organized into groups of three to work collaboratively on a one-hour creative writing task. The purpose of the task was to promote teamwork and to assess the effectiveness of group collaboration in different conditions. To control for experimental variables, we turned off the chatbot feature in conditions 3 and 4 to exclude its effect on teamwork goal setting. Instead, we provided participants in these two conditions with goal-setting guidance based on the SMART framework to help them clarify their collaborative goals and task requirements. We collected each group's creative writing output through the TeamSynerPad system and collected each group's engagement and collaboration quality scores utilizing a questionnaire.

Afterward, quantitative analysis evaluates the collaborative effectiveness, engagement, and creative performance under different conditions.

Through this experiment, we expect to understand the specific effects of different functions in the TeamSynerPad system on teamwork and creative writing and how to optimize the system design to promote the team's creative performance and collaborative efficiency. The experiment results will provide valuable insights for further developing the TeamSynerPad system and empirical evidence for studying teamwork and technology-supported creative activities.

4.3. Participants

Because this study aimed to simulate collaboration among student teams in higher education, we recruited participants with at least a bachelor's degree background in education.

To estimate the required sample size, we conducted a power analysis using GPower. In this between-groups design-based experiment, we had four different conditions and set up four groups in the GPower analysis. We set an alpha error probability of 0.05 and a statistical efficacy of 0.8 to obtain a large effect size f of 0.4. This effect size was chosen based on the results of previous relevant studies and our expectations of the experimental effects.

Based on the results of the G*Power calculations, we arrived at a total sample size of 76. However, because the experimental design called for participants to engage in creative writing in subgroups of three, we needed to ensure that the number of participants in each condition was divisible by three. To accomplish this, we decided to recruit a total of 84 people so that there would be 21 people in each condition, divided into 7 subgroups, and result in 28 subgroups throughout the experiment.

Thus, our experiment will consist of four conditions with seven subgroups of three people in each condition. This organizational structure meets the experimental design requirements and allows us to compare results across conditions effectively.

4.4. Procedure

The experiment consisted of four parts, as shown in FIGURE 4.2. First, we recruited 84 participants with higher education backgrounds and assigned them to 28 groups (three people per group) in a randomized manner. Subsequently, we randomly assigned these 28 groups to the four conditions described in the experiment.

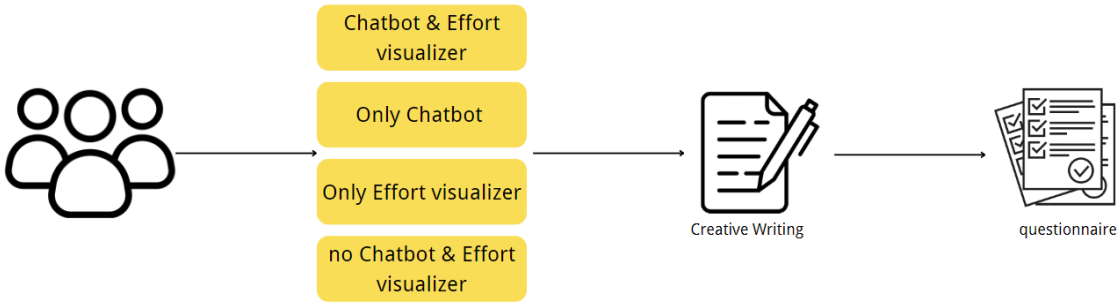


Figure 4.2: Experiment procedure flow

Before the experiment starts, all participants must attend a training phase to familiarize themselves with the operation of the TeamSynerPad system, including its visualization components and how to use the chatbot. The training was designed to ensure that all participants could utilize the system's functionality effectively for the ensuing creative writing task. After the training session, participants were informed of the background and intent of the experiment, and after signing an informed consent form, the experiment officially started.

The core part of the experiment was the creative writing task, in which each group must collaborate to complete a creative writing story within 1 hour. The start time of the experiment was informed to all participants, and the supervisor would notify participants of the end of the time after 1 hour. The supervisor monitored this session to ensure that all participants adhered to the rules and time constraints of the experiment.

Upon completing the creative writing task, participants were required to complete an online questionnaire to measure user engagement and quality of collaboration. The questionnaire was administered through the Google Forms platform to gather participant feedback on the collaborative process and the experience of using the system. Finally, creative writing outputs from each group will be collected to assess the quality of creative writing.

4.5. Measures

User Engagement

In digital collaboration environments, user engagement is an important indicator of whether team members are engaged and actively participating in collaborative tasks. It also serves as an indicator of the overall efficiency of team collaboration. High user engagement is usually associated with better teamwork and higher quality of creative output. Therefore, using a valid and reliable tool to measure user engagement is essential to understand the effectiveness of team collaboration under different conditions.

We employed the User Engagement Scale (UES) Short Form [34] because it is a validated and widely used tool for measuring user engagement in digital environments. Its short form is particularly suitable for our experimental setup, allowing for fast and efficient data collection without overburdening the participants. The UES short form employs a scale range of 1-5, where 1 represents strong disagreement and 5 represents strong agreement. Its robustness and reliability ensure that we can trust the data collected to draw reliable conclusions.

Applying the UES Short Form in the post-experiment questionnaire will enable us to collect data on user engagement. This data will be compared with the quality of the team's creative writing and the effectiveness of the collaboration, thereby gaining a deeper understanding of the effectiveness of teamwork under different conditions.

Collaboration Effectiveness

To assess the collaborative effectiveness of teams under different conditions, we utilized the Comprehensive Assessment of Team Member Effectiveness - Likert's Easy Version (CATME) [35]. The CATME is a tool designed for self-and peer-assessment within a team [16], primarily used to assess the performance of team members in various aspects of team dynamics. It is widely implemented in educational and professional settings to enhance team effectiveness and personal accountability.

The CATME focuses on five main areas to understand the extent of team members' contributions:

- **Contributing to the Team's Work:** Fulfilling responsibilities, coming prepared to team meetings, and making significant contributions.
- **Interacting with Teammates:** Effective communication, timely information exchange, and providing feedback.
- **Keeping the Team on Track:** Assessing progress, staying aware of external factors, and helping in planning and organizing work.
- **Expecting Quality:** Aiming for success and high-quality output from the team.
- **Relevant Knowledge, Skills, and Abilities (KSAs):** Possessing necessary skills and expertise for the job and understanding other team members' roles.

The CATME scale, assessed on a 5-point Likert scale from 1 (Strongly Disagree) to 5 (Strongly Agree), allows us to quantitatively analyze teamwork efficacy while keeping the assessment process concise.

Its widespread use in education and the rigorous design and validation process ensure its reliability and validity in assessing teamwork efficacy. Implementing the CATME will provide quantitative data on teamwork efficacy for our study and help us understand the impact of different technological tools on teamwork and efficacy. This understanding will inform the optimization of the TeamSynerPad system to enhance teamwork and collaboration effectiveness.

By evaluating collaboration efficacy under different conditions quantitatively, we aim to gain insights into the impact of various features in the TeamSynerPad system on team collaboration effectiveness. Concurrently, comparing and analyzing this data with the team's creative writing quality and user engagement allows us to comprehensively assess the overall effectiveness of team collaboration under different scenarios.

Creative Writing Quality

In our research, we employed the PaperRater platform [33] to assess the creative writing quality produced by each group participating in the experiment. PaperRater [33] is a web-based writing assistance and assessment tool, offering instant grammar checking, spelling correction, plagiarism detection, and advice on writing style and structure. It utilizes natural language processing techniques and artificial

intelligence algorithms to analyze and assess textual content quality. PaperRater provides quick feedback, identifies potential errors, and suggests improvements. Beyond basic grammar and spelling checks, it features readability analysis, vocabulary diversity assessment, and overall writing quality scoring. In scoring the creative writing results, we chose PaperRater for its objective and consistent evaluation criteria, ensuring uniform evaluation across all works, crucial for the validity of the experiment. PaperRater's ability to check grammatical and spelling errors and evaluate style, structure, and readability is particularly important in assessing creative writing quality. Its rapid feedback aids in promptly understanding and improving writing quality. As an online platform, PaperRater is easily accessible and user-friendly, requiring no complicated setup. Additionally, its detailed feedback not only aids in

4.6. Data Processing

In the data processing phase of this study, we start by focusing on four different scenarios included in the experimental design:

- The use of both chatbot and effort visualizer
- The use of chatbot only
- The use of effort visualizer only
- The non-use of these tools at all

The experiment involved 28 groups, seven in each case, each consisting of three members, for a total of 84 participants. Upon completing the experimental tasks, participants were asked to complete two questionnaires. The first was the User Engagement Scale (UES) Short Form, used to assess user engagement. The second was the Comprehensive Assessment of Team Member Effectiveness - Likert's Easy Version (CATME), which assessed team member effectiveness. Both questionnaires use a rating scale of 1 to 5. During data processing, we first calculated the average score for each group. Specifically, we averaged the scores of all members within each group on the UES and CATME questionnaires, respectively, to represent the User Engagement score and Effectiveness score for that group. The following data analysis plan compares the differences in group scores across experimental situations to assess the impact of chatbot, effort visualizer, and their combinations on user engagement and team effectiveness.

5 Results

This chapter discusses the experimental results of this study in-depth and is organized into three main sections. First, in Section 5.1, we conducted a detailed analysis of the data collected under different experimental conditions through descriptive statistics to reveal the underlying trends and characteristics of the data. Subsequently, in Section 5.2, we conducted a series of statistical tests to validate the hypotheses presented in Chapter 4 and assess the impact of chatbots and effort visualizers on collaboration effectiveness and user engagement. Finally, in Section 5.3, we summarized the differences between chatbot and effort visualizer in improving collaboration effectiveness.

5.1. Descriptive Statistics

In this study, we conducted detailed descriptive statistical analyses of user engagement and collaboration effectiveness to reveal trends and characteristics of the data across experimental conditions. Each experimental condition consisted of seven groups, and we collected a total of 84 observations and calculated the average user engagement and collaboration effectiveness to represent the feedback score of this group based on three people in each group.

5.1.1. User Engagement

Table 5.1 shows the results of the descriptive statistics of user engagement under different experimental conditions. When both the chatbot (✓) and the effort visualizer (✓) were used, user engagement had the highest mean value of 3.793 and a standard deviation of 0.220. This indicates that user engagement was higher and relatively stable in this condition. In contrast, when the chatbot and effort visualizer (✗, ✗) were not used, the mean of user engagement was the lowest at 3.120 with a standard deviation of 0.159, suggesting that user engagement was lower and less variable in this condition.

The box plot 5.1a further reveals the distribution of user engagement in different conditions. We can observe that the median and interquartile range of user engagement are higher when chatbots are used (with or without effort visualizers), suggesting that chatbots may play a positive role in increasing user engagement. While there were some outliers, they did not appear to have a significant impact on the overall trend.

Chatbot	Effort Visualizer	N=28	Mean	SD	Min	25%	50%	75%	Max
✓	✓	7	3.79	0.22	3.47	3.67	3.72	3.99	4.06
✓	✗	7	3.63	0.15	3.39	3.57	3.64	3.72	3.83
✗	✓	7	3.40	0.16	3.17	3.31	3.42	3.50	3.61
✗	✗	7	3.12	0.16	3.00	3.02	3.03	3.18	3.42

Table 5.1: Descriptive statistics for user engagement under different conditions

5.1.2. Collaboration Effectiveness

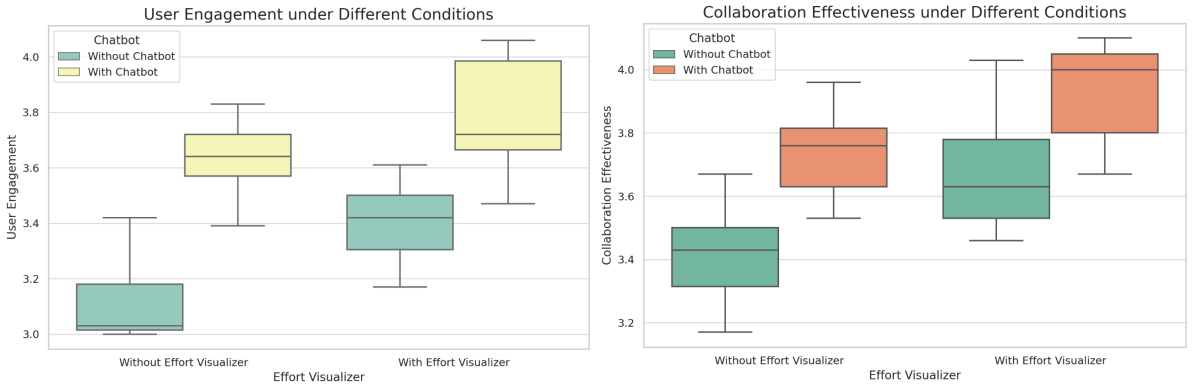
Table 5.2 shows the results of descriptive statistics of collaboration effectiveness under different experimental conditions. Similar to the results for user engagement, the highest mean value of collaboration

effectiveness, 3.924, with a standard deviation of 0.175, was found when both the chatbot and the effort visualizer were used, suggesting that the combination of the chatbot and the effort visualizer may have a positive effect on improving collaboration effectiveness. In the condition without the chatbot and effort visualizer, the collaboration efficiency had the lowest mean of 3.414 and a standard deviation of 0.171.

The boxplot of collaboration effectiveness shows that the distribution of collaboration efficiency is relatively centralized across all conditions, but the median and interquartile range of collaboration efficiency are higher in the conditions where chatbots are used (✓, ✓ and ✓, ✗). This further suggests that chatbots may have a positive impact on improving collaboration efficiency.

Chatbot	Effort Visualizer	N=28	Mean	SD	Min	25%	50%	75%	Max
✓	✓	7	3.92	0.18	3.67	3.80	4.00	4.05	4.10
✓	✗	7	3.73	0.16	3.53	3.63	3.76	3.82	3.96
✗	✓	7	3.68	0.20	3.46	3.53	3.63	3.78	4.03
✗	✗	7	3.41	0.17	3.17	3.32	3.43	3.50	3.67

Table 5.2: Descriptive statistics for collaboration effectiveness under different conditions

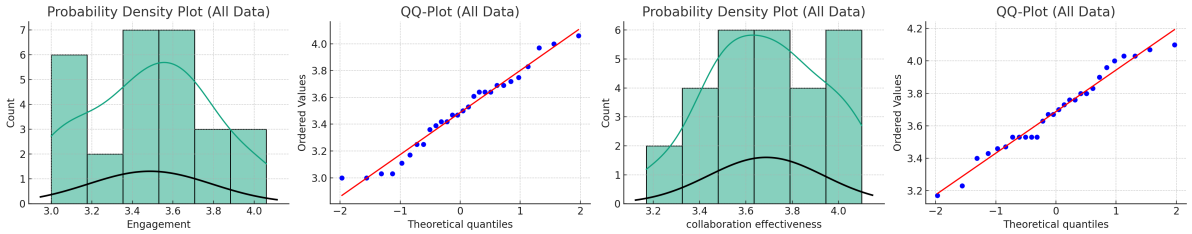


(a) User Engagement

(b) Collaboration Effectiveness

Figure 5.1: Box plot of user engagement and collaboration effectiveness under Different Conditions

Through descriptive statistics and boxplot analyses of user engagement and collaboration effectiveness, we found that the conditions of using chatbots and effort visualizers showed high levels and stability of both user engagement and collaboration effectiveness. These findings provide an initial perspective and foundation for the statistical testing and hypothesis validation that we conduct in the subsequent sections.



(a) User Engagement probability density and QQ-plot

(b) Collaboration Effectiveness probability density and QQ-plot

Figure 5.2: Probability density and QQ-plot of user engagement and collaboration effectiveness

Following the descriptive statistical analysis and visualization of the data, our next step is to conduct a two-way ANOVA test. This will allow us to rigorously test our hypotheses and assess the impact of using a chatbot and effort visualizer on user engagement and collaboration effectiveness under different conditions. Before proceeding with the ANOVA test, it is crucial to verify the assumption of

normality for the data, as this is a prerequisite for conducting parametric tests. To this end, we utilized the Kolmogorov-Smirnov (K-S) test, a non-parametric method, to evaluate whether our samples are likely to be drawn from a normal distribution. The results of the K-S test suggest that the data for user engagement and collaboration effectiveness are similar to a normal distribution. This finding aligns with the visual inspections we performed earlier, as depicted in Figure 5.2, where the probability density function plots and Q-Q plots demonstrate the normal distribution characteristics of our data. These results support the validity of proceeding with the two-way ANOVA test in our subsequent analysis.

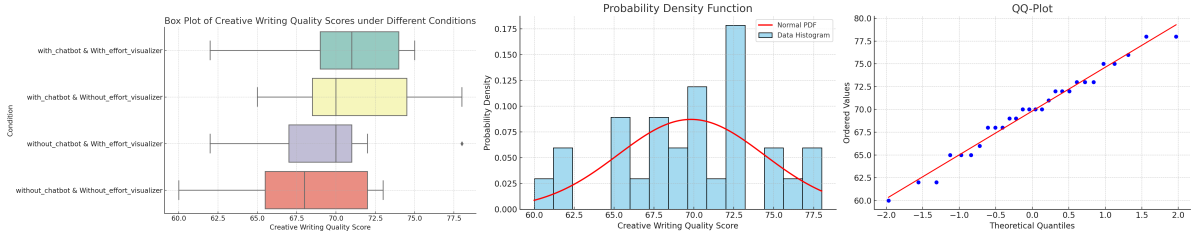
5.1.3. Creative Writing Result

Table 5.3 shows the results of descriptive statistics of collaboration effectiveness under different experimental conditions. When both the chatbot (✓) and the effort visualizer (✓) are used, the mean of the creative writing results is 70.57 with a standard deviation of 4.58. When the chatbot and the effort visualizer (✗, ✗) are not used, the mean of the creative writing results is 68, with a standard deviation of 4.73. When only the chatbot (✓) was used without the effort visualizer (✗), the mean of the creative writing results was 71.29, with a standard deviation of 4.61. When only the effort visualizer (✓) was used without the chatbot (✗), the mean of the creative writing results was 71.29, with a standard deviation of 4.61. From these results, it is clear that with the different features of TeamSynerPad, the results of the creative writing were similar.

With the boxplot 5.3a results shown in Figure 5.4a, we can observe that the distribution of creative writing quality is relatively concentrated in all conditions, with no significant outliers. The median and quartile spacing are also relatively close across conditions, suggesting that the effects of the chatbot and effort visualizer on creative writing quality may not be particularly significant.

Chatbot	Effort Visualizer	N	Mean	SD	Min	25%	50%	75%	Max
✓	✓	7	70.57	4.58	62	69	71	74	75
✓	✗	7	71.29	4.61	65	68.5	70	74.5	78
✗	✓	7	69.43	5.09	62	67	70	71	78
✗	✗	7	68.00	4.73	60	65.5	68	72	73

Table 5.3: Descriptive statistics for creative writing quality under different conditions



(a) Creative Writing Result Box Plot

(b) Creative Writing Result probability density function

Figure 5.3: Box plot of user engagement and collaboration effectiveness under Different Conditions

Similarly, to verify our hypothesis of the effect of using the chatbot and effort visualizer on the quality of creative writing, we need to conduct a two-way ANOVA test. We also used the Kolmogorov-Smirnov (K-S) test to verify that the data on the quality of creative writing follows a normal distribution. The results of the K-S test indicate that the users' creative writing quality is normally distributed. Figure 5.3b shows the probability density function plot and Q-Q plot of creative writing quality, which is consistent with the results of the K-S test.

5.2. Hypothesis Test

To address our hypothesis and ensure a rigorous evaluation, we set the significance level (α) at 0.05. The specific conditions pertaining to each experimental group are detailed in figure 4.1.

5.2.1. Effect of TeamSynerPad On User Engagement

H1a: *Using a chatbot positively impacts user engagement.*

To analyze the main effect of the chatbot on user engagement in a collaborative setting, we conducted a two-way ANOVA test. The results in Table 5.3 showed a significant difference ($p<0.001$) between user engagement with and without the chatbot, controlling for the use or non-use of the effort visualizer. To further validate this finding, a Tukey's HSD test was conducted. The test results confirmed this significant difference and showed that the mean of user engagement for the chatbot group was 0.46 higher than that of the group that did not. This difference was statistically significant ($p=0.0015$). The results indicate that at the alpha = 0.05 significance level, chatbot significantly affects user engagement. Therefore, we accept hypothesis H1a.

H1b: *Using the effort visualizer positively impacts user engagement.*

similarly, to analyze the main effect of the effort visualizer on user engagement in a collaborative setting. The two-way ANOVA test result in Table 5.3 showed a significant difference ($p = 0.026$) in user engagement. To further validate this finding, we proposed a Tukey's HSD test. The test results shows p-value = 0.0528. The results of Tukey's HSD test show that the difference in average user engagement between using the effort visualizer and not using the effort visualizer is not statistically significant. The results suggest that although we have some evidence to support that the effort visualizer affects user engagement, this effect may not be as strong as we have observed through ANOVA. We can accept hypothesis H1b, but more data or further research may be needed to confirm the exact impact of the effort visualizer on user engagement.

H1c: *The use of both chatbot and effort visualizer has a more significant impact on user engagement than using only one tool.*

Finally, in order to analyze the effect of the simultaneous use of chatbot and escort visualizer on user engagement, the results of the Two-way ANOVA test in Table 5.3 showed that there was no significant interaction effect between chatbot and effort visualizer ($p = 0.356$). This result suggests that using a combination of the two does not have a significantly different effect on user engagement than using only one of them. Further analysis revealed that Tukey's HSD test exhibited the same result ($p = 0.337$), further confirming that the combined use of chatbot and effort visualizer did not exhibit higher user engagement. Therefore, we reject H1c.

Factor	Sum of Squares	Degrees of Freedom	F-Value	P-Value
Chatbot	1.44	1	48.05	<0.0001
Effort Visualizer	0.34	1	11.34	0.003
Chatbot:Effort Visualizer	0.026	1	0.88	0.36
Residual	0.72	24	-	-

Table 5.4: Two-way ANOVA results for the effect of Chatbot and Effort visualizer on User Engagement.

To summarize, our hypothesis testing results indicate that using a Chatbot has a significant positive impact on user engagement, while using an Effort Visualizer also tends to enhance user engagement, though its impact is not statistically significant compared to that of the Chatbot. Furthermore, although the experimental group that used both the Chatbot and Effort Visualizer simultaneously showed the best performance in our sample, the two-way ANOVA test did not provide clear evidence to conclude that their combined use has a significantly higher impact on user engagement than using either one alone. This suggests that the synergistic effect of using both tools together is not significantly greater than the effect of using them individually.

Comparison	Mean Difference	Adjusted P-Value	Significant
Chatbot & Effort Visualizer vs. Chatbot	0.16	0.34	No
Chatbot & Effort Visualizer vs. Effort Visualizer	0.39	0.002	Yes
Chatbot & Effort Visualizer vs. None	0.67	0.001	Yes
Chatbot vs. Effort Visualizer	0.23	0.08	No
Chatbot vs. None	0.51	0.001	Yes
Effort Visualizer vs. None	0.28	0.03	Yes

Table 5.5: Tukey's HSD Test Results for User Engagement

5.2.2. Effect of TeamSynerPad On Collaboration Effectiveness

H2a: *Using a chatbot positively impacts user engagement.*

In exploring the role of chatbots in collaborative tasks, we found that it had a significant positive impact on teams' collaboration effectiveness. A two-way ANOVA showed that teams using chatbots had significantly higher collaboration effectiveness scores compared to teams not using chatbots (p=0.000280). This result was further confirmed by Tukey's HSD test, which showed that the group using chatbots exhibited higher collaboration effectiveness compared to the group that did not use the tool, and the difference was statistically significant (p=0.0121). Therefore, there is good reason to accept hypothesis H2a.

H2b: *Using an effort visualizer positively impacts user engagement.*

The effort visualization tool also significantly affected collaboration effectiveness (p=0.0024). However, Tukey's HSD test showed that although some of the between-group comparisons showed significant differences, this effect was relatively small and less significant than the effect of the chatbot (p=0.047). The results suggest that effort visualization tools have a positive effect on teamwork, but their impact may not be as significant as chatbots. Therefore, we accept hypothesis H2b.

H2c: *The use of both chatbot and effort visualizer has a more significant impact on collaboration effectiveness than using only one tool.*

Finally, the two-way ANOVA test did not find a significant interaction effect between the chatbot and the effort visualization tool (p=0.590). The results of Tukey's HSD test also showed that the combined use of the two did not result in an additional collaborative effect enhancement. Therefore, we believe that the use of both the chatbot and the effort visualization tool did not lead to a more significant collaboration effectiveness enhancement than the use of either of them alone, thus rejecting hypothesis H2c.

Factor	Sum of Squares	Degrees of Freedom	F-Value	P-Value
Chatbot	0.57	1	18.05	0.0003
Effort Visualizer	0.36	1	11.51	0.0024
Chatbot:Effort Visualizer	0.010	1	0.30	0.59
Residual	0.75	24	-	-

Table 5.6: Two-way ANOVA results for the effect of Chatbot and Effort Visualizer on Score.

To summarize, our hypothesis testing results suggest that using a Chatbot significantly enhances collaboration effectiveness. At the same time, using an Effort Visualizer also appears to have a positive impact, though its effect is not statistically significant compared to the Chatbot. Furthermore, even though the group that utilized both the Chatbot and Effort Visualizer demonstrated the highest performance in our sample, the two-way ANOVA test did not provide conclusive evidence to assert that their combined use significantly outperforms either tool alone. The result implies that the synergistic effect of using both tools together does not significantly exceed the impact of using them individually on collaboration effectiveness.

Comparison	Mean Difference	Adjusted P-Value	Significant
Chatbot & Effort Visualizer vs. Chatbot	0.19	0.21	No
Chatbot & Effort Visualizer vs. Effort Visualizer	0.25	0.07	No
Chatbot & Effort Visualizer vs. None	0.51	0.001	Yes
Chatbot vs. Effort Visualizer	0.06	0.90	No
Chatbot vs. None	0.32	0.012	Yes
Effort Visualizer vs. None	0.26	0.05	Yes

Table 5.7: Tukey's HSD Test Results for Collaboration Effectiveness

5.2.3. Effect of TeamSynerPad On Creative Writing Result

H3a: *Using a chatbot positively impacts Creative Writing Quality.*

Table 5.8 presents the results of the two-way ANOVA test. We found that using a chatbot does not have a significant impact on creative writing quality, as indicated by a p-value of 0.229, which is greater than the significance level of 0.05. Therefore, we have enough evidence to reject H3a, and we conclude that using a chatbot does not significantly impact creative writing quality.

H3b: *Using the effort visualizer positively impacts Creative Writing Quality.*

Similarly, as demonstrated in Table 5.8. We found that using an effort visualizer does not have a significant impact on creative writing quality, as indicated by a p-value of 0.844, which is greater than the significance level of 0.05. Thus, we reject H3b.

H3c: *The use of both chatbot and effort visualizer has a more significant impact on user engagement than using only one tool.*

Furthermore, the interaction effect between the chatbot and effort visualizer on creative writing quality was also examined. The two-way ANOVA results revealed a p-value of 0.56 for the interaction effect, which is greater than the significance level of 0.05. The result suggests that there is no significant difference in creative writing quality between using both the chatbot and effort visualizer together and using them individually. Therefor, we do not find support for H3c, indicating that the combined use of a chatbot and effort visualizer does not lead to a significantly greater impact on creative writing quality. Thus, we reject H3c.

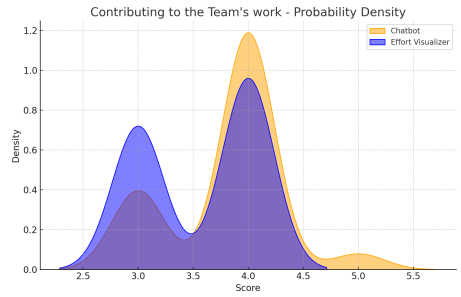
Factor	Sum of Squares	Degrees of Freedom	F-Value	P-Value
Chatbot	34.32	1	1.52	0.23
Effort Visualizer	0.89	1	0.040	0.84
Chatbot:Effort Visualizer	8.04	1	0.36	0.56
Residual	542.86	24	-	-

Table 5.8: Two-way ANOVA results for the effect of Chatbot and Effort Visualizer on creative writing quality.

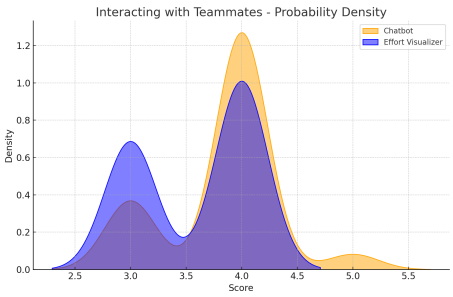
5.3. Exploratory Findings

The results of the two-way ANOVA showed that using both the chatbot and effort visualizer together did not result in further improvements in user engagement and collaboration effectiveness compared to using only one of the tools. Thus, we decided to conduct a more in-depth comparative analysis of using chatbot alone and report visualizer alone to explore the differences in facilitating collaboration. As we presented in section 4, the CATME scale assesses collaborative effectiveness on five dimensions: ”Contribution to teamwork,” ”Interaction with teammates,” ”Keeping the team on schedule,” ”Quality of expectations,” and ”Possession of relevant knowledge and skills.” Therefore, we plotted probability density distributions and performed independent sample t-tests using the chatbot and effort visualizer.

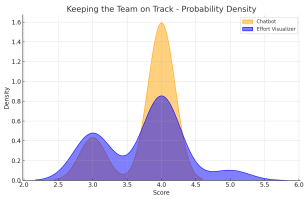
As shown in figure 5.4. On the dimension of Contributing to the Team's work, the distribution of chatbot shows a higher concentration trend and kurtosis, indicating that most participants gave higher ratings. This may indicate that chatbot performs better in helping team members contribute to the team's work. In the dimension of Interacting with Teammates, chatbot similarly showed a more concentrated and higher distribution of ratings than effort visualizer. This suggests that chatbot is more



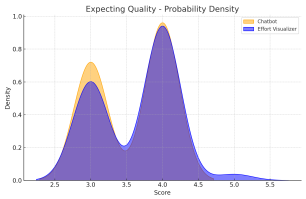
(a) Contributing to the Team's work



(b) Interacting with Teammates



(c) Keeping the Team on Track



(d) Expecting Quality

(e) Having relevant knowledge and skills

Figure 5.4: Probability Density of Scores for Chatbot vs Effort Visualizer

popular in facilitating interactions among team members. For the dimensions Keeping the Team on Track, Expecting Quality, and Having relevant knowledge and skills, the distributions of the two conditions were close to each other, with no significant difference, indicating that there was no significant difference in the performance of the two tools in these three dimensions.

The t-test results further validate our conclusion. As shown in table 5.9, at a significant level of alpha = 0.05, there is a statistically significant difference between the use of chatbot compared to the use of the Effort Visualizer on the dimension of ”contribution to teamwork” with a p-value of 0.020561, which is smaller than 0.05. It shows that in this dimension, chatbot performs better compared to effort visualizer.

Similarly, as shown in table 5.9, at a significant level of alpha = 0.05, there is a statistically significant difference of 0.020561 between the use of chatbot compared to the use of the Effort Visualizer on the dimension of ”Interacting with Teammates”. visualizer, there is a statistically significant difference of 0.0298 p-value which is smaller than 0.05. The results indicate that chatbot is more able to facilitate team member communication in this dimension as compared to the effort visualizer.

Finally, as shown in table 5.9, at a significant level of alpha = 0.05, the p-values for the dimensions ”Keeping the Team on Track,” ”Expecting Quality,” and ”Having relevant knowledge and skills,” the p-values for using chatbot compared to using the Effort Visualizer were 0.669, 0.529, and 0.113, respectively, which are all greater than 0.05. The results show no significant difference in the three dimensions of using the chatbot compared to using the Effort visualizer.

In reviewing the results of the two-way analysis of variance (ANOVA) and Tukey's Honest Significant Difference (HSD) test in Section 5.2.3, we noted that the use of both the chatbot and the effort visualizer alone improved the effectiveness of teamwork compared to the situation without the aid of the tools. However, no significant difference was found between the two tools when they were compared. Overall, the chatbot and effort visualizer have similar effectiveness in facilitating collaboration. Continuing to examine the results of the t-tests on the five dimensions of collaboration effectiveness, at a significant level of alpha = 0.05, we found that although the overall collaboration effectiveness was comparable, the chatbot was more effective on the dimensions of ”Contribution to teamwork” and ”Interaction with teammates” than effort visualizer. ”This finding emphasizes the potential of chatbots to facilitate more effective communication styles in collaborative environments.

Collaboration level	T-Statistic	P-Value	Chatbot Std. Dev.	Visualizer Std. Dev.
Contributing to the Team's work	2.17	0.03	0.51	0.50
Interacting with Teammates	2.21	0.03	0.49	0.50
Keeping the Team on Track	0.430	0.67	0.42	0.59
Expecting Quality	-0.63	0.528521	0.50	0.53
Having relevant knowledge and skills	-1.60	0.11	0.50	0.35

Table 5.9: T-Test Results for Chatbot and Effort Visualizer Conditions

6 Discussions

6.1. Key Findings

The purpose of this study is to explore how technology can be used to facilitate student collaboration and, in turn, enhance student engagement and the effectiveness of teamwork. Given the increased emphasis on the importance of collaborative learning in the current educational landscape, this study looks at developing and evaluating technological tools that can support this educational goal. In particular, we focus on the use of two technologies, chatbots and effort visualizers, to explore how they work separately and together in the collaborative process of students.

To address RQ 1 **How can we design an interactive system to promote student collaboration?**

Through an in-depth literature review, we recognized that good communication, clearly set collaboration goals, ensuring goal alignment, and enhancing team transparency are key factors in promoting effective collaboration. This study further found that chatbots are effective in improving team communication and, when combined with the SMART framework, can also assist in setting collaborative goals. Meanwhile, applying effort visualization techniques can significantly increase transparency within the team, making each member's contribution more visible. Based on these insights, we designed a system called *TeamSynerPad*, which combines a Chatbot for setting collaboration goals and an Effort Visualizer for increasing team transparency to improve user engagement and collaboration effectiveness.

To address RQ 2 **How do different features of the collaborative system influence user engagement and collaboration effectiveness?**

We conducted an experimental test on a creative writing task to evaluate the system we designed to facilitate student collaboration. As described in Chapter 4, we designed a between-group experimental design with four conditions:

1. Use of both a chatbot and an effort visualizer,
2. Use of only a chatbot,
3. Use of only an effort visualizer,
4. A control group that did not use any tools.

Participants were randomly assigned to these conditions and completed the creative writing task in the corresponding situations.

Upon completion of the experiment, we collected data on participants' engagement and collaborative effects through a questionnaire. Then, we proposed a series of statistical tests to validate our hypotheses and came up with the following finding: both the chatbot and the effort visualizer, used alone, significantly increased student engagement and collaboration effects. However, when the two techniques were used in combination, no additional boosting effect was observed. This finding suggests that while each tool was effective individually, their combination did not produce the expected synergistic effects. Additionally, this study found that the effects of using a chatbot and effort visualizer

to improve user engagement were approximately the same, with users who used a chatbot compared to an effort visualizer demonstrating more efficient communication in a collaborative task setting.

This finding provides insight into how we can promote student collaboration through technology interventions and provides an empirical basis for future educational technology design and instructional strategies. It also points to directions for future research, such as exploring how different types of collaboration tools can be used more effectively in combination and how they can be implemented in different instructional settings.

6.2. Limitations

Although this study provides valuable insights into the impact of student collaboration, we must acknowledge several limitations that may affect the broad applicability of our findings.

First, the sample size of the experiment was limited. There were a total of 84 participants in this study, which provides a degree of statistical power for analysis. When conducting the G*power sample size prediction, we chose a large effect size f of 0.4, which was to be able to observe a significant difference. Typically, a medium effect size f is usually set to 0.25 in order to observe appropriate differences. If we had set the effect size to 0.25, we would have needed more participants to achieve equal statistical efficacy. Due to resource constraints, we could not expand our sample size, which may have limited the generalizability of our findings and the explanatory power of our results.

Second, the research design was a short-term experiment in which participants were divided into groups of three and completed a creative writing task with a time limit of one hour. However, real-life student-team projects typically require weeks or months of ongoing collaboration. Therefore, this experiment cannot capture the possible impact of the long-term use of chatbots and effort visualization tools on teamwork effectiveness. The dynamics on a long-term scale may differ from the results of our short-term experiment.

Finally, the nature of the task is also a limitation. We chose a creative writing task to simulate the process of students writing reports and thinking together. While this helps to simulate some types of academic collaboration, actual teamwork tasks are more complex and involve knowledge and skills from more specialized fields. Our participants came from various professional backgrounds, which may have influenced their behaviors and interactions in the collaboration. In real-world team tasks, students need to deal with more complex division of labor and collaborative goal setting, factors that are not adequately captured in the current experimental design.

As such, these limitations highlight the boundaries of current research findings in terms of their explanatory power and applicability. While our findings provide initial guidance on the design and application of technology tools in student collaboration, they also reveal areas that need to be explored in depth by future research. In the next sections, we discuss in detail how these limitations can be addressed in future work and propose new research directions based on the results of this study.

6.3. Future Work

Based on this study, we proposed the following recommendations for future study.

For future research, one area worth exploring is developing a generalized conversational framework to support teams in setting collaborative goals across disciplines and diverse collaborative tasks. Such a framework should go beyond the constraints of a single academic task and provide customized collaboration strategies based on the specific needs and team dynamics of various projects. Also, considering the great potential of large-scale language models such as ChatGPT in education, future research could include how these models can facilitate team communication, answer questions, and provide real-time project management support.

For the effort visualization tool, the design of our effort visualizer is based on arol's research, which determines the percentage of words students contribute in real time based on the number of words they contribute. However, we recognized the challenges of rationally visualizing everyone's contributions when dealing with complex collaborative tasks. Therefore, we need to develop a method to efficiently assign task weights and make the team division of tasks more transparent.

Additionally, we recognize the need for long-term research to explore the impact of the continued use of chatbots and effort visualization tools on student collaborative behavior. While short-term experiments can provide observations of immediate effects, long-term studies will help us understand the long-term effects of these tools on students' collaborative habits, motivation, and efficiency during

ongoing use.

Finally, although this study showed that combining different collaboration aids did not significantly improve user engagement and collaboration efficiency, we can still see their potential when used alone. Therefore, it is hoped that future research will lead to the development of an integrated system that will combine the strengths of the different tools to provide a more efficient and intuitive solution for team project management and collaboration. Such an integrated system would include adapting and optimizing the functionality of existing tools, as well as developing new features to facilitate efficient collaboration between students from different backgrounds and majors.

7

Conclusion

In this study, we used an empirical approach to explore the impact of a system that incorporates a chatbot with goal-setting functionality and an efficacy visualizer demonstrating team effort on user engagement and collaboration effectiveness in a student collaboration environment. We used a between-group design with four scenarios involving 84 participants, we reproduced collaboration patterns among students in a creative writing task, and we observed the effects of chatbots and efficacy visualizers on students' collaborative effectiveness in different scenarios of use.

Overall, the experimental results support the idea that both the chatbot and the effectiveness visualization tool, when used individually, positively impacted team collaboration efficiency. The use of both also promoted higher levels of engagement among participants. Although the data showed that the chatbot and the effectiveness visualization tool appeared to perform best when used together, the experimental results needed to be sufficiently robust to establish that using the two in combination would have a stacking effect to enhance collaboration efficiency. The experimental result suggests that the effects of the two tools may be independent of each other. Furthermore, when comparing the effects of chatbots and effort visualization tools on team collaboration efficiency, we found that although they performed similarly in improving overall efficiency, chatbots showed a more significant effect in facilitating team communication.

These results reveal the potential of chatbots and efficacy visualization tools in stimulating student collaboration and highlight their unique value in promoting user engagement and communication efficiency. While our study provides initial supportive evidence for using these tools in educational technology, further research is needed to explore their effectiveness in different educational contexts, tasks, and when used over time. Future work should focus on optimizing the design of these tools to maximize their collaborative benefits, as well as further research on how to integrate these tools into students' learning lives to enhance the overall learning experience.

References

[1] Noorhan Abbas et al. “Online chat and chatbots to enhance mature student engagement in higher education”. In: *International Journal of Lifelong Education* 41.3 (2022), pp. 308–326.

[2] Oluseyi Oluseun Adesina et al. “Managing group work: the impact of peer assessment on student engagement”. In: *Accounting Education* 32.1 (2023), pp. 90–113.

[3] Rohan Ahuja et al. “A Digital Dashboard for Supporting Online Student Teamwork”. In: *Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing*. 2019, pp. 132–136.

[4] A Aranzabal, E Epelde, and M Artetxe. “Team formation on the basis of Belbin’s roles to enhance students’ performance in project based learning”. In: *Education for Chemical Engineers* 38 (2022), pp. 22–37.

[5] FX Risang Baskara. “Chatbots and Flipped Learning: Enhancing Student Engagement and Learning Outcomes through Personalised Support and Collaboration”. In: *IJORER: International Journal of Recent Educational Research* 4.2 (2023), pp. 223–238.

[6] Sanam Shirazi Beheshitha et al. “The role of achievement goal orientations when studying effect of learning analytics visualizations”. In: *Proceedings of the sixth international conference on learning analytics & knowledge*. 2016, pp. 54–63.

[7] Ivo Benke, Michael Thomas Knierim, and Alexander Maedche. “Chatbot-based emotion management for distributed teams: A participatory design study”. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW2 (2020), pp. 1–30.

[8] May Britt Bjerke and Ralph Renger. “Being smart about writing SMART objectives”. In: *Evaluation and program planning* 61 (2017), pp. 125–127.

[9] Antonela Čižmešija, Ana Horvat, and D Plantak Vukovac. “Improving student engagement and course completion using Chatbot application”. In: *INTED2021 Proceedings*. IATED. 2021, pp. 8346–8354.

[10] Helen Donelan and Karen Kear. “Online group projects in higher education: persistent challenges and implications for practice”. In: *Journal of Computing in Higher Education* (2023), pp. 1–34.

[11] George T Doran et al. “There’sa SMART way to write management’s goals and objectives”. In: *Management review* 70.11 (1981), pp. 35–36.

[12] James E Driskell, Eduardo Salas, and Tripp Driskell. “Foundations of teamwork and collaboration.” In: *American Psychologist* 73.4 (2018), p. 334.

[13] Jiahui Du, Weijiao Huang, and Khe Foon Hew. “Supporting students goal setting process using chatbot: implementation in a fully online course”. In: *2021 IEEE International Conference on Engineering, Technology & Education (TALE)*. IEEE. 2021, pp. 35–41.

[14] Andreea Gheorghe et al. “The effect of humor and perceived social interdependence on teamwork engagement in student groups”. In: *Current Psychology* 42.27 (2023), pp. 23830–23841.

[15] Ella Glikson et al. “Visualized automatic feedback in virtual teams”. In: *Frontiers in psychology* 10 (2019), p. 814.

[16] Chiriac E Hammar. “Group work as an incentive for learning–students’ experiences of group work”. In: *Frontiers in psychology* 5.558 (2014), pp. 1–10.

[17] Khe Foon Hew et al. “Using chatbots to support student goal setting and social presence in fully online activities: learner engagement and perceptions”. In: *Journal of Computing in Higher Education* 35.1 (2023), pp. 40–68.

[18] Weijiao Huang et al. “Using chatbot as a goal-setting facilitator in an online course: students’ performance and teacher’s perception”. In: *2022 IEEE International Conference on Teaching, Assessment and Learning for Engineering (TALE)*. IEEE. 2022, pp. 389–395.

[19] Petra Isenberg et al. “Collaborative visualization: Definition, challenges, and research agenda”. In: *Information Visualization* 10.4 (2011), pp. 310–326.

[20] Mohammad Khalil and Magnus Rambech. “Eduino: A telegram learning-based platform and chatbot in higher education”. In: *International Conference on Human-Computer Interaction*. Springer. 2022, pp. 188–204.

[21] Alexandra Kolm et al. “International online collaboration competencies in higher education students: A systematic review”. In: *Journal of Studies in International Education* 26.2 (2022), pp. 183–201.

[22] Mohammad Amin Kuhail et al. “Interacting with educational chatbots: A systematic review”. In: *Education and Information Technologies* 28.1 (2023), pp. 973–1018.

[23] Jeya Amantha Kumar. “Educational chatbots for project-based learning: investigating learning outcomes for a team-based design course”. In: *International journal of educational technology in higher education* 18.1 (2021), pp. 1–28.

[24] K Blaine Lawlor. “Smart goals: How the application of smart goals can contribute to achievement of student learning outcomes”. In: *Developments in business simulation and experiential learning: Proceedings of the annual ABSEL conference*. Vol. 39. 2012.

[25] Hoi Yan Lin and Jia You. “Predicting teamwork performance in collaborative project-based learning.” In: *Journal of Education and Learning* 10.4 (2021), pp. 104–117.

[26] Edwin A Locke. “Motivation through conscious goal setting”. In: *Applied and preventive psychology* 5.2 (1996), pp. 117–124.

[27] Edwin A Locke and Gary P Latham. “Building a practically useful theory of goal setting and task motivation: A 35-year odyssey.” In: *American psychologist* 57.9 (2002), p. 705.

[28] Edwin A Locke and Gary P Latham. “Goal setting theory: The current state”. In: *New developments in goal setting and task performance*. Routledge, 2013, pp. 623–630.

[29] Pilar Martín-Hernández et al. “The Moderating Role of Teamwork Engagement and Teambuilding on the Effect of Teamwork Competence as a Predictor of Innovation Behaviors among University Students”. In: *International Journal of Environmental Research and Public Health* 19.19 (2022), p. 12047.

[30] Jade McKay and Bhavani Sridharan. “Student perceptions of collaborative group work (CGW) in higher education”. In: *Studies in Higher Education* (2023), pp. 1–14.

[31] Jennifer D Nahrgang et al. “Goal setting in teams: The impact of learning and performance goals on process and performance”. In: *Organizational Behavior and Human Decision Processes* 122.1 (2013), pp. 12–21.

[32] Rupayan Neogy, Jonathan Zong, and Arvind Satyanarayan. “Representing real-time multi-user collaboration in visualizations”. In: *2020 IEEE Visualization Conference (VIS)*. IEEE. 2020, pp. 146–150.

[33] Vi Thong Nguyen. “Automated essay assessment: an evaluation on paperrater’s reliability from practice/Nguyen Vi Thong”. In: *Journal of Creative Practices in Language Learning and Teaching (CPLT)* 5.1 (2017), pp. 1–18.

[34] Heather L O’Brien, Paul Cairns, and Mark Hall. “A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form”. In: *International Journal of Human-Computer Studies* 112 (2018), pp. 28–39.

[35] Matthew W Ohland and Misty L Loughry. “Designing a Peer Evaluation Instrument that is Simple, Reliable, and Valid: The Comprehensive Assessment of Team-Member Effectiveness”. In: *STEM Assessment Conference*. 2006, p. 202.

[36] Vitaliy Popov et al. “Multicultural student group work in higher education: An explorative case study on challenges as perceived by students”. In: *International Journal of Intercultural Relations* 36.2 (2012), pp. 302–317.

[37] Elena De Prada, Mercedes Mareque, and Margarita Pino-Juste. “Teamwork skills in higher education: is university training contributing to their mastery?” In: *Psicologia: Reflexao e Critica* 35 (2022).

[38] Nitin Rane. “Chatbot-Enhanced Teaching and Learning: Implementation Strategies, Challenges, and the Role of ChatGPT in Education”. In: *Challenges, and the Role of ChatGPT in Education (July 21, 2023)* (2023).

[39] Amon Rapp et al. “Collaborating with a Text-Based Chatbot: An Exploration of Real-World Collaboration Strategies Enacted during Human-Chatbot Interactions”. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–17.

[40] Donghoon Shin et al. “IntroBot: Exploring the Use of Chatbot-assisted Familiarization in Online Collaborative Groups”. In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 2023, pp. 1–13.

[41] Mateus da Silveira Colissi et al. “A Chatbot that Uses a Multi-agent Organization to Support Collaborative Learning”. In: *International Conference on Human-Computer Interaction*. Springer. 2021, pp. 31–38.

[42] Sofie Smedegaard Skov et al. “Designing a conversational agent to promote teamwork and collaborative practices using design thinking: An explorative study on user experiences”. In: *Frontiers in Psychology* 13 (2022), p. 903715.

[43] Melissa Sundberg, Raina Paul, and George R Verghese. “Teamwork and collaboration”. In: *Patient Safety and Quality in Pediatric Hematology/Oncology and Stem Cell Transplantation* (2017), pp. 51–65.

[44] Freddy Tapia et al. “A Container Orchestration Development that Optimizes the Etherpad Collaborative Editing Tool through a Novel Management System”. In: *Electronics* 9.5 (2020), p. 828.

[45] Amanuel G Tekleab, Narda R Quigley, and Paul E Tesluk. “A longitudinal study of team conflict, conflict management, cohesion, and team effectiveness”. In: *Group & organization management* 34.2 (2009), pp. 170–205.

[46] Leigh Thompson. “Improving the creativity of organizational work groups”. In: *Academy of Management Perspectives* 17.1 (2003), pp. 96–109.

[47] Sandeep A Thorat and Vishakha Jadhav. “A review on implementation issues of rule-based chatbot systems”. In: *Proceedings of the international conference on innovative computing & communications (ICICC)*. 2020.

[48] Tri Chau Minh Tri. “Creating smart, human-like chatbot for businesses using BotPress platform”. PhD thesis. 2020.

[49] Hung Wei Tseng and Hsin-Te Yeh. “Team members’ perceptions of online teamwork learning experiences and building teamwork trust: A qualitative study”. In: *Computers & Education* 63 (2013), pp. 1–9.

[50] Sebastian Wollny et al. “Are we there yet?-a systematic literature review on chatbots in education”. In: *Frontiers in artificial intelligence* 4 (2021), p. 654924.

[51] Yan Xiao et al. “Building high performance surgical teams: Theoretical and practical considerations for team consistency”. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 58. 1. SAGE Publications Sage CA: Los Angeles, CA. 2014, pp. 748–752.

[52] Shunan Zhang et al. “I know my teammates: the role of Group Member Familiarity in Computer-Supported and face-to-face collaborative learning”. In: *Education and Information Technologies* (2023), pp. 1–17.

[53] Xi Zhang et al. “Learning analytics in collaborative learning supported by Slack: From the perspective of engagement”. In: *Computers in Human Behavior* 92 (2019), pp. 625–633.

A

User Engagement Scale

Please rate each item by circling the number that best represents your level of agreement with the following statements using the scale provided:

Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
1	2	3	4	5

1. I am engaged myself in this experience.
2. The time I spent using the Application TeamSynerPad just slipped away.
3. I was absorbed in this experience.
4. I felt frustrated while using this Application TeamSynerPad.
5. I found this Application TeamSynerPad confusing to use.
6. Using this Application TeamSynerPad was taxing.
7. This Application TeamSynerPad was attractive.
8. This Application TeamSynerPad was aesthetically appealing.
9. This Application TeamSynerPad appealed to my senses.
10. Using Application TeamSynerPad was worthwhile.
11. My experience was rewarding.
12. I felt interested in this experience.

B

CATME Scale

APPENDIX B

Comprehensive Assessment of Team Member Effectiveness—Behaviorally Anchored Rating Scale (BARS) Version

	Your name					← Write the names of the people on your team including your own name. This self and peer evaluation asks about how you and each of your teammates contributed to the team during the time period you are evaluating. For each way of contributing, please read the behaviors that describe a "1", "3," and "5" rating. Then confidentially rate yourself and your teammates.
Contributing to the Team's Work	5	5	5	5	5	• Does more or higher-quality work than expected. • Makes important contributions that improve the team's work. • Helps to complete the work of teammates who are having difficulty.
	4	4	4	4	4	Demonstrates behaviors described in both 3 and 5.
	3	3	3	3	3	• Completes a fair share of the team's work with acceptable quality. • Keeps commitments and completes assignments on time. • Fills in for teammates when it is easy or important.
	2	2	2	2	2	Demonstrates behaviors described in both 1 and 3.
	1	1	1	1	1	• Does not do a fair share of the team's work. Delivers sloppy or incomplete work. • Misses deadlines. Is late, unprepared, or absent for team meetings. • Does not assist teammates. Quits if the work becomes difficult.
Interacting with Teammates	5	5	5	5	5	• Asks for and shows an interest in teammates' ideas and contributions. • Improves communication among teammates. Provides encouragement or enthusiasm to the team. • Asks teammates for feedback and uses their suggestions to improve.
	4	4	4	4	4	Demonstrates behaviors described in both 3 and 5.
	3	3	3	3	3	• Listens to teammates and respects their contributions. • Communicates clearly. Shares information with teammates. Participates fully in team activities. • Respects and responds to feedback from teammates.
	2	2	2	2	2	Demonstrates behaviors described in both 1 and 3.
	1	1	1	1	1	• Interrupts, ignores, bosses, or makes fun of teammates. • Takes actions that affect teammates without their input. Does not share information. • Complains, makes excuses, or does not interact with teammates. Accepts no help or advice.
Keeping the Team on Track	5	5	5	5	5	• Watches conditions affecting the team and monitors the team's progress. • Makes sure that teammates are making appropriate progress. • Gives teammates specific, timely, and constructive feedback.
	4	4	4	4	4	Demonstrates behaviors described in both 3 and 5.
	3	3	3	3	3	• Notices changes that influence the team's success. • Knows what everyone on the team should be doing and notices problems. • Alerts teammates or suggests solutions when the team's success is threatened.
	2	2	2	2	2	Demonstrates behaviors described in both 1 and 3.
	1	1	1	1	1	• Is unaware of whether the team is meeting its goals. • Does not pay attention to teammates' progress. • Avoids discussing team problems, even when they are obvious.
Expecting Quality	5	5	5	5	5	• Motivates the team to do excellent work. • Cares that the team does outstanding work, even if there is no additional reward. • Believes that the team can do excellent work.
	4	4	4	4	4	Demonstrates behaviors described in both 3 and 5.
	3	3	3	3	3	• Encourages the team to do good work that meets all requirements. • Wants the team to perform well enough to earn all available rewards. • Believes that the team can fully meet its responsibilities.
	2	2	2	2	2	Demonstrates behaviors described in both 1 and 3.
	1	1	1	1	1	• Satisfied even if the team does not meet assigned standards. • Wants the team to avoid work, even if it hurts the team. • Doubts that the team can meet its requirements.
Having Relevant Knowledge, Skills, and Abilities	5	5	5	5	5	• Demonstrates the knowledge, skills, and abilities to do excellent work. • Acquires new knowledge or skills to improve the team's performance. • Able to perform the role of any team member if necessary.
	4	4	4	4	4	Demonstrates behaviors described in both 3 and 5.
	3	3	3	3	3	• Has sufficient knowledge, skills, and abilities to contribute to the team's work. • Acquires knowledge or skills needed to meet requirements. • Able to perform some of the tasks normally done by other team members.
	2	2	2	2	2	Demonstrates behaviors described in both 1 and 3.
	1	1	1	1	1	• Missing basic qualifications needed to be a member of the team. • Unable or unwilling to develop knowledge or skills to contribute to the team. • Unable to perform any of the duties of other team members.

Figure B.1: CATME scale

C

Informed Consent Form

Promoting student collaboration through technology interventions

Before agreeing to participate in the study, it is important that you read the following explanation (pages 1 and 2). You (a participant) will be given a copy of this Informed Consent Form.

Title of Research

Promoting student collaboration through technology interventions. This research is conducted in the context of a Master's thesis.

Researchers

- Shikuan Li, Master student of Delft University of Technology, Computer Science
- Ujwal Gadiraju, Assistant Professor, Web Information Systems group of the Faculty of Engineering, Mathematics and Computer Science, Delft University of Technology, NL
- Tahir Abbas, Postdoctoral researcher, Web Information Systems group of the Faculty of Engineering, Mathematics and Computer Science, Delft University of Technology, NL

The team members are henceforth referred to as "Researchers."

Purpose of the Research

The purpose of the study is to investigate the effectiveness of combining chatbots with visualization techniques in facilitating collaborative goal setting.

Confidentiality and Data Management

During our research on the effectiveness of combining chatbots with visualization techniques in facilitating collaborative goal setting in a creative writing task, we will collect data related to your interactions with the chatbot and your writing results for the creative writing task as well as the questionnaire result.

Anonymity: All your inputs and interactions will be recorded anonymously. Participants will be labeled with codes such as "user1" and "user2", ensuring that the data collected cannot be traced back to an individual as well as ensuring non-demographic data will be collected. All the inputs and interactions will be kept up to 6 months after the completion of the study and will be deleted afterward. The anonymized data may be used for research and publications.

For the project duration (until October 2023), the experiment data will be stored on the SURFdrive cloud storage. The data will be available to the researchers in this project with organizational login access.

Explanation of Procedures

In this experiment, you will be randomly matched with two other people in a team—a total group of three to complete a creative writing task. The creative writing task will occur in four different scenarios: chatbot with visual dashboard, chatbot only, visual dashboard only, and no other technical support. You will be randomly assigned to one of the four scenarios.

The experiment is expected to last two hours. Within the first half hour, you and your group will be shown which scenario to complete your creative writing under and how to use our online collaboration platform. You and your group members will complete the creative writing task in the next hour. In the last half hour, you will participate in our survey to complete questionnaires.

Risks

Participation in this creative writing task is considered low-risk. All involvement is anonymous, and participants' identities within the study will be concealed, labeled as "user1", "user2", and so on. Thus, personal information will not be disclosed.

While we do not anticipate this task to cause discomfort to the participants, writing may sometimes touch upon personal experiences and emotions. As such, specific prompts may elicit mild emotional distress in some individuals. If you feel discomfort about the creative writing task or decide you no longer wish to continue at any point, you have the right to cease participation.

Withdrawal without Prejudice

Taking part in the study is voluntary. You or your guardian are free to withdraw consent, discontinue the study, or require the retrieval of your data at any time, without prejudice from the researchers. Should you wish to receive a record of your data generated through the experiment, please inform the researchers. They will provide you with a copy after the experiment is finished.

Questions

Any questions concerning this study and/or in the case of issues due to the research can be directed to the researcher, Shikuan Li, either in person or via S.Li-56@student.tudelft.nl.