



**HAL**  
open science

# Impact of communication modalities on social presence and regulation processes in a collaborative game

Anthony Basille, Élise Lavoué, Audrey Serna

► **To cite this version:**

Anthony Basille, Élise Lavoué, Audrey Serna. Impact of communication modalities on social presence and regulation processes in a collaborative game. *Journal on Multimodal User Interfaces*, 2025, 10.1007/s12193-024-00450-z . hal-04939639

**HAL Id: hal-04939639**

**<https://hal.science/hal-04939639v1>**

Submitted on 11 Feb 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Impact of communication modalities on social presence and regulation processes in a collaborative game

Anthony Basille<sup>1</sup>[0000-0002-2733-2107], Élise Lavoué<sup>1</sup>[0000-0002-2659-6231], and Audrey Serna<sup>2</sup>[0000-0003-1468-9761]

- <sup>1</sup> University of Lyon, University Jean Moulin Lyon 3, iaelyon school of Management, CNRS, LIRIS, UMR5205, F-69621 Lyon, France  
{anthony.basille, elise.lavoue}@univ-lyon3.fr
- <sup>2</sup> Univ Lyon, INSA Lyon, CNRS, UCBL, LIRIS, UMR5205, F-69621 Villeurbanne, France  
audrey.serna@insa-lyon.fr

**Abstract.** In the digital era, leveraging communication technologies to foster collaborative learning is of utmost importance. This study explores the impact of different communication modalities, such as text, audio and video, on social presence and regulation processes within a computer-supported collaborative learning (CSCL) environment. Using learning analytics, we examine the influences of these modalities on collaboration and derive recommendations for their optimized use in the design of future CSCL environments. Our findings reveal a significant impact of communication modalities on the sense of social presence and regulation of collaborative activities. Audio communication results in enhanced co-presence, psychobehavioral accessibility, and better regulation processes compared to video and text modalities, indicating that audio is the most suitable modality in collaborative virtual environments for decision-making tasks. Conversely, video communication still facilitated strategic planning and enhanced self-regulation. Chat communication showed the lowest sense of social presence, yet improvements over time suggest that participants adapt to this modality, enhancing their collaborative efficiency.

**Keywords:** collaboration · regulation · social presence · communication modalities

## 1 INTRODUCTION

Also called “21st-century skills”, the 4Cs – critical thinking, collaboration, communication, and creativity – have been widely recognized as crucial for education and work [49, 20]. Defined as the act of two or more people working towards a common goal [21], the importance of collaboration, among these 4 skills, is increasingly recognized in learning, as learning becomes more social and interactive.

In parallel, technology has evolved to offer a range of environments for collaborative learning, moving beyond traditional classrooms. Computer-supported collaborative learning (CSCL) environments and collaborative virtual environments (CVEs), for instance, provide multiple communication channels, such as chat, audio, video, and immersive virtual reality (VR), playing a key role in collaboration processes [12]. Some studies have demonstrated that technological features related to communication can influence perceptions of users, in particular the sense of social presence [56], and can shape social behaviors [24].

Social presence refers to the sensation of being with others in a mediated environment [29]. This contributes to collaboration, community building, and user satisfaction in virtual spaces [80]. Regulation involves planning, monitoring, evaluating, and adjusting shared activities in a collaborative setting [37]. Working together implies the co-construction of representations of shared tasks, objectives and strategies [37]. Social presence and regulation are therefore complementary in a collaborative environment, giving isolated people the impression of being together in the mediated environment and enabling them to work together more effectively. However, they are still studied separately and we believe that there is a lack of understanding of how communication features can influence social presence and regulation and shape collaboration.

We conducted a study that aims to explore to what extent different communication modalities (text, audio, video) have an impact on collaboration, more particularly on social presence and regulation processes in a CSCL setting. To achieve this aim, we developed a collaborative game involving 3 players. We ran a study with 72 participants across 3 conditions, varying communication modalities, each condition comprising 8 groups of 3 players.

In this paper, we first review existing research on collaborative learning environments, regulation processes and social presence in such environments. We then describe our study and the way we used learning analytics to examine regulation processes through the lens of communication modalities, in relation with the sense of social presence felt by the users. We then discuss our findings and draw implications for integrating these communication modalities to enhance the design of future CSCL environments. The aim is to contribute to the understanding of the role of technology in enhancing collaborative learning experiences and outcomes, providing insights for educators, researchers, and practitioners in the field.

## 2 RELATED WORK

### 2.1 Collaborative learning environments

The dynamics of collaboration, particularly in CSCL and CVEs, have garnered increasing interest. These technologies have been recognized as important resources to enhance cognitive performance [48, 35], stimulate knowledge construction [68], and ultimately improve learning outcomes.

Studies corroborate that students engaged in CSCL environments report higher levels of learning [30], contribute to more comprehensive reports, par-

ticipate more equally in the learning process [23, 34], and are involved in more complex and stimulating discussions than when working alone [38]. Additionally, students often report greater satisfaction in CSCL settings compared to traditional learning groups [23].

However, alongside these promising findings, there are also challenges and less positive experiences reported within CSCL environments. Some students find discussions more confusing [73], less productive [69, 70], and more time-consuming [23] as compared to traditional face-to-face settings. In addition, students in these environments sometimes report lower participation rates [52], increased conflict [31], reduced group cohesiveness [70], and decreased satisfaction [5]. These mixed experiences can be attributed to both the design of the environments and the social and cognitive behaviors of the groups [50].

In CSCL and CVEs, communication channels play a central role in the facilitation of effective interactions [26]. Varied modalities, from text and audio to video and virtual reality, offer unique attributes influencing collaborative tasks efficacy [33]. For example, video communication, in contrast to text, transmits non-verbal cues like facial expressions and gestures, crucial for understanding and reducing miscommunication [17]. Yet, these advanced modalities also present challenges, such as potential for misinterpretation and feelings of disconnect. Some students in CSCL environments have reported experiences of confusion [73], reduced participation [52], increased conflict [31], and decreased satisfaction [5]. While technology offers the potential to mimic or even enhance face-to-face interactions, the nature of virtual collaboration introduces a set of unique challenges that must be understood and navigated judiciously.

Not every collaborative task requires the depth offered by high-richness modalities. Sometimes, simpler text or audio chat might suffice due to their accessibility and ease of use. Therefore, while selecting the appropriate communication modality, it is crucial to consider the nature of the collaborative task, the technological proficiency of the participants, and the desired outcomes [19, 18, 14].

In studying CSCL environments, we turn our attention to the critical role of regulation processes in fostering successful collaboration, emphasized by Järvelä et al. [39] as an essential element in resolving the communication difficulties encountered in these environments. Emerging from both individual and group dynamics, these processes serve as the foundation upon which learning in CSCL is built. Our study primarily focuses on delving deep into these regulatory mechanisms, given their substantial influence in shaping collaborative experiences and enhancing the effectiveness of learning outcomes.

## 2.2 Regulation processes

Regulation represents the processes that members use to plan, monitor, evaluate and regulate the joint activity. In a collaborative activity, members need to regulate their behavior, their cognitive and emotional states to progress in the game and achieve the learning objectives [75]. It is widely accepted that effective learners use a combination of cognitive, behavioral, and motivational strategies to self-regulate and reinforce their learning during academic pursuits

[63]. Considering the collaborative nature of group learning, recent research has emphasized the integration of individual and social processes in the study of interpersonal regulation, recognizing the central role of social context in students' self-regulated learning [42, 28]. This recognition of the importance of social context in self-regulated learning has led to the emergence of theories and models of self-regulation in highly interactive and dynamic learning situations, namely co-regulation and shared regulation [27].

Self-regulation, defined by Zimmerman [81], is an individualized, learner-centric process that involves self-generated thoughts, feelings, and behaviors aimed at achieving specific goals. It includes individuals' self-regulated learning strategies that are reflexively responsive to feedback, contributing to learners' ability to modify their actions, motivation, and understanding of a subject matter. This transformative mechanism essentially turns inherent mental abilities into functional academic skills, aligning them with learners' personal objectives.

Co-regulation, as depicted by Hadwin et al. [28], involves an interactive, shared problem-solving process between participants that can contribute to the development of individual learning and reasoning abilities. In a collaborative environment, each member can potentially influence the regulation process of another, contributing to both individual and collective progression toward a shared goal. This interactive influence often materializes through inquiries or feedback on actions, reasoning, goals, or future strategies.

Finally, shared social regulation, based on Hadwin et al. [28] and Järvelä et al. [37], seeks to instigate self-regulation within a group context. This form of regulation involves collective construction of objectives and rules, focusing specifically on setting, monitoring and evaluating group goals and interactions.

Only limited research has been conducted yet on how communication modalities affect regulation processes. Several studies have shown that the online reduction of social cues (such as chat communication) and the asynchronicity or semi-synchronicity of interactions can cause individuals to feel anonymous and less self-aware, ultimately decreasing their ability to self-regulate and see themselves from others' perspectives [15, 43, 79]. To compensate for the limitations of the communication medium and the absence of social cues, participants must adjust their language [74, 78]. This adaptation is not immediate and can potentially affect regulation processes.

In recent studies, Järvelä et al. [40] underline the challenges related to the observation and understanding of regulation processes. They underline the importance of considering the situated and cyclical nature of these processes. The authors argue that empirical research should target not only the occurrence of regulatory activities but also how the regulation of learning changes over time. In addition, they insist on the lack of methods to "make visible" mental processes that are often invisible and the limitation of subjective measures (such as video coding). If self-regulation can be assessed with data logs analysis, this limitation is particularly true for shared social regulation, which implies analyzing deeply group conversations as done in Järvelä et al. [41].

In this study, we are thus interested in observing how communication modalities influence regulation processes and their evolution over time. We propose to adopt an objective approach using data logs to analyze regulation processes, focusing more specifically on self-regulated and co-regulated processes.

### 2.3 Social presence

Presence, in this context, is described as the subjective experience of being in the virtual environment and is grounded on the ability to do things in the virtual environment, relying on the actions and the affordances offered in the virtual environment. Usually, presence is divided into different sub-concepts: spatial, self and social presence [51]. The latter is particularly interesting for technology-mediated communication.

Indeed, social presence is defined as the feeling of "being with others" [29]. It is the moment-to-moment awareness of the co-presence of a mediated body and the sense of accessibility of the other beings' psychological, emotional, and intentional states [10]. This can be interpreted as the degree to which a person is perceived as "real" in mediated communication.

Exploring the social dimensions of perceived presence is important, particularly in CSCL settings. The notions of togetherness, co-presence, and social presence address social interaction and collaboration issues. The different works on social presence have shown that it has an impact on many factors, for instance allowing to create a comfortable social climate, to create a productive collaboration, or to increase participation and interaction [53, 72, 1, 80].

We rely on the definition of Biocca and Harms [10], who distinguishes three dimensions: 1) co-presence, defined as the degree to which the users feel that they are not alone and isolated. Goffman et al. [25] extends this notion by incorporating "awareness of others", meaning that the user is aware of the mediated other, and the other is aware of the user; 2) psychobehavioral accessibility, which focuses on the perceived accessibility of the other, the users' sense of awareness and access to the other (through attentional engagement, emotional state or understanding); 3) perceived symmetry, defined as the degree of symmetry or correlation between the users' sense of social presence and their perception of their partner's sense of social presence [10]. This dimension relies on the assertion that social presence is not only based on our vision of the other but also on our feeling of what the other thinks of us.

Historically, the level of social presence in virtual environments was viewed largely through a technocentric lens, with media richness theory positing that some mediums inherently offer richer communication due to their technological capabilities [17]. In essence, richer technologies were seen as offering a deeper sense of "being there." Contrarily, Walther's social information processing theory (SIPT) [77] introduced a more human-centric approach. It suggests that given adequate time, individuals can adapt to any medium, whether rich or limited in cues, to achieve meaningful communication. This adaptability underscores the significance of individual communication strategies. The hyperpersonal model further expands on this by suggesting that computer-mediated communication

(CMC) can occasionally yield even greater intimacy than face-to-face interactions due to selective self-presentation [78].

However, while these theories emphasize human adaptability, the inherent features of a communication modality cannot be entirely dismissed. We believe that they may influence social presence, especially in constrained scenarios. In essence, while technological affordances play a role in shaping social presence, human adaptability and strategic communication also hold a significant way. For effective virtual interactions, it is essential to consider both technological potential and the agency of users. To support this statement, we can draw on existing literature. This includes studies on the impact of the level of interactivity with a virtual agent on the sense of social presence [66], the quality of audio with a 6-channel (Dolby 5.1) or 2-channel (Dolby Stereo) audio system [67], and the presence of haptic feedback when the action is performed with a partner [46]. Additionally, studies have been conducted to compare communication modalities. A systematic review conducted by Oh et al. [56] indicates that while audio and video modalities may have a comparative advantage over text for communication, the strength of audiovisual modalities over audio-only modalities is more nuanced, and the results are not necessarily significant. To underscore the fact that linear increases in immersion do not necessarily lead to corresponding increases in social presence, Oh et al. [56] suggest that one of the main differences may be the nature of the task proposed by the studies.

In summary, this section highlights the multifaceted nature of social presence in virtual environments and emphasizes its significance. The literature underscores that social presence is not only a product of technological features but is also significantly shaped by human adaptability and communication strategies. This study aims to expand on previous research by investigating how different communication modalities affect the sense of social presence and, consequently, impact regulation processes over time. An objective approach will be adopted using a social presence questionnaire to analyze social presence.

### 3 Research questions

Given the theoretical landscape presented in previous sections, our research aims to study to what extent different communication modalities impact collaboration, specifically focusing on the dimensions of social presence and regulation processes within a computer-supported cooperative learning (CSCL) context. This research contributes to our understanding of how these modalities can be optimized in the design and implementation of CSCL environments, and how they shape collaboration, social presence, and regulation processes. We more specifically address the following research questions:

- RQ 1: What are the impacts of communication modalities on social presence? We aim to investigate the differential effects of various communication modalities (text, audio, video) on social presence perceived by the users. We further intend to explore the sub-dimensions of co-presence: psychobe-

havioral accessibility, and perceived symmetry under different modalities to provide a deeper understanding of the impact on social presence.

- RQ 2: What are the impacts of communication modalities on regulation processes? We analyze the impact of different communication modalities on the regulation processes, especially self-regulation and co-regulation.
- RQ 3: How do regulation processes evolve across different phases of collaborative activities? We aim to examine the evolution of regulation processes between different parts of collaborative activities. The intent is to uncover if and how the evolution of regulation processes is influenced by the selected communication modality.
- RQ 4: Is there a correlation between social presence and regulation processes, and does this correlation vary across different communication modalities?

## 4 Description of the collaborative serious game

We developed a collaborative turn-based serious game, whose gameplay involves a collaborative effort among three players with the end goal of repairing four breaches within a submarine. As the game progresses, the water level within the submarine rises gradually, culminating in-game loss once it exceeds a predetermined threshold. Therefore, the players’ task is to complete the repairs before the submarine is entirely submerged.

In a single turn, each player can take several actions. They are allotted one movement point and two action points (they can remove two levels of water, or repair one breakdown). The repair of a breakdown requires the concurrent presence of at least two players in the same room. Each player has a distinct power facilitating the execution of specific actions. For instance, the diver has an additional movement point per turn, the oceanographer can drain the rooms adjacent to his, and the computer scientist can transfer water from one room to another each turn.

In this study, players were given the ability to interact with each other at any point. Communication modalities, depending on the experimental conditions, could involve video and audio streams with chat, solely audio and chat, or exclusively chat. Outside their turn, players viewed the game from a third-person perspective as shown in figure 1. During their turn, the viewpoint switched to first-person 2, and players had to undertake their actions directly within the room. Information such as individual room water level, the submarine blueprint, the total water level in the submarine, and the remaining time was not disclosed to the players playing (in first person viewpoint).

The choice of information asymmetry in the design of the game is based on the findings from one of our previous study [6]. We found that when switching from third- to first-person view, limiting access to global information (like water levels and game map) creates a natural need for communication and coordination between players. The collaborative task was specifically designed as a time-constrained, decision-making game requiring continuous coordination. Players



must make strategic decisions about resource allocation (movement and scooping), coordinate repairs that require multiple players' presence, and manage time pressure from rising water levels.

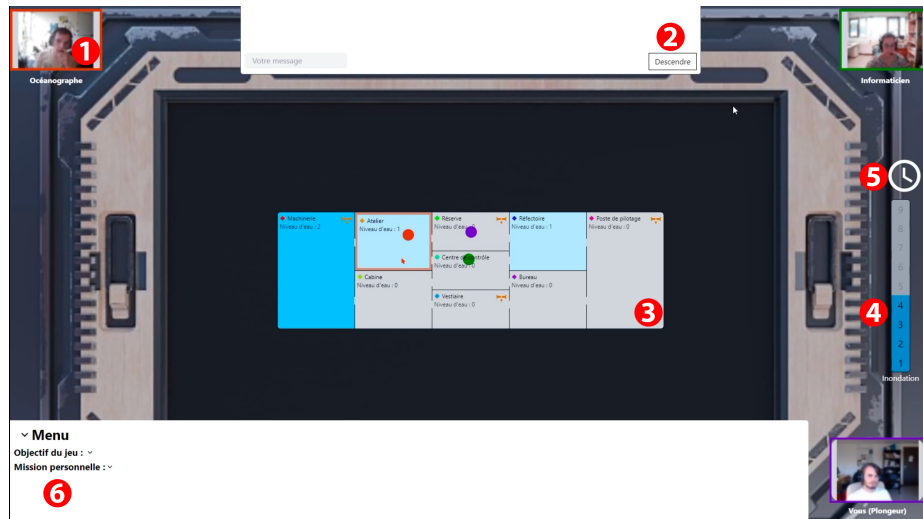


Fig. 1: **Game Interface Used in the Study, Third-Person perspective.** 1) Players' webcam for video communication, 2) Chat window for player communication, 3) Submarine room layout, 4) Current water levels indicator (in the submarine), 5) Clickable timer displaying remaining time, 6) Game menu with main goal and sub-goal.



Fig. 2: **Game Interface Used in the Study, First-Person perspective.** 1) Players' webcam for video communication, 2) Chat window for player communication, 3) Clickable flashlight for environment exploration and item discovery, 4) Game menu with goal, sub-goal, and action options, 5) Highlighted pathway indicating possible movement in the room, 6) Submarine breach requiring repair.

In alignment with the complex dynamics of collaborative learning environments, Collaborative Learning Analytics (CLA) emerges as a vital tool in navigating the wealth of data generated during collaborative learning sessions [54], and can make sense of it for both learners and teachers [55, 76].

Evaluating collaborative activities necessitates gauging collaboration quality. Depending on measurement objectives, various indicators can be analyzed. Role determination might involve examining topic clusters [16, 59]. Dominance can be inferred from talk time [47, 3, 9, 58], and turn-taking frequency can indicate participation levels [45]. Leveraging these metrics, stakeholders can better understand and improve collaborative learning.

In our study, we focus on regulation processes, and more specifically on 16 indicators derived from the literature [71], as well as the definitions of regulation processes presented in section 2.2. These indicators are divided into 3 categories: task-related indicators, self-regulation indicators, and co-regulation indicators. Task-related indicators are not linked to the self-regulation or co-regulation processes but are the main actions required to complete the game. To ensure clarity and comprehensibility, we have only included the most relevant indicators and those that yielded significant results.

We identified the following task-related indicators:

- SpecialPowerUses: indicates the number of times a special power is used during the game. Special powers are unique abilities provided to the players to

facilitate task completion. This indicator could reveal players' strategic use of resources to overcome rising water levels, repair breaches and an understanding of the game and the means at their disposal.

- TurnsWithoutScooping: represents the number of turns the player has not performed the water scooping action.

The following indicators are categorized under self-regulation as they exemplify the individual's autonomous engagement in goal-directed activities, strategic planning, and personal resource management, in alignment with Zimmerman's concept of self-regulated learning and behavior.

- OwnObjectTaken: indicates the number of times a personal object is retrieved in the game. Each player's item has been hidden in a different room of the submarine, and all players could see them. The retrieval of personal objects can be linked to individual strategies or learning methods to manage resources and navigate the game. It reflects how players regulate their behavior to achieve personal goals, balancing between individual objectives and collective needs.
- ObjectsTaken: denotes the number of times other players' objects are stolen. This could be indicative of strategic behaviors or competitive instincts. It could also affect the player's focus on an individual mission rather than the collective one. If the item is stolen, the player who has been robbed will not be able to complete their personal mission. This indicator could reflect competitive instincts and individual strategic planning in pursuit of personal goals.
- FlashlightUseSum: total time spent using the flashlight. As the flashlight is what enables the player to find objects in a room, it indicates the time allocated to carry out a personal mission. It could be representative of an individual desire to explore the environment, or a feeling of curiosity to better understand it, to achieve personal goals.
- ShowSubGoal: number of times a player clicked to display a sub-goal. This indicator is linked to one's own personal mission, and therefore the quest for success from an individual viewpoint. A high frequency might indicate a need for clarity, underlining self-reflection and planning as part of the self-regulation processes.

The following indicators fall under the category of co-regulation, reflecting the collaborative and interactive aspects of learning and problem-solving within a group, in line with the description of co-regulation processes and collective goal attainment proposed by Hadwin et al. [27]

- TurnsPlayed: refers to the number of turns played by the players. Higher number of turns can reflect a higher level of active engagement and shared problem-solving, which is key in co-regulation. Given that the more the game progresses, the more difficult it becomes to win (the number of water levels rises faster and faster), it is essential to try to win with as few turns as possible.

- `CommonObjectTaken`: measures the number of times the common object is retrieved. The retrieval of common objects might denote collaborative efforts and shared responsibility toward achieving the secondary collective goal. This indicates a collaborative effort towards a shared goal, resonating with the concept of interactive, shared problem-solving in co-regulation.
- `MeanTurn`: refers to the average duration of a turn. Shorter turns might signify a quick pace of action and decision-making in the team.
- `CancelledAction`: number of actions canceled during the game. When an action is performed, a pop-up appears, and the player can either confirm or cancel the action. Frequent cancellations may signal conflicts, misunderstandings, or changes in strategies, indicating less efficient co-regulation within the team. Frequent cancellations might indicate ongoing negotiation, strategy adjustments, or conflict resolution, key aspects of co-regulation in a team setting.
- `ImpossibleAction`: reflects the number of actions deemed impossible by the players. It is possible to perform an action, but if conditions or rules don't allow it, it won't succeed (no more movement points, for example). This can lead to misunderstandings of the rules of the game, which can result from less efficient communication among team members.
- `ShowMenu`: indicates the number of times a player clicks to display the menu (bring up menu to display actions, access to the ability to display main goal and sub-goal)
- `ShowMainGoal`: represents the number of times a player clicks to display the main goal. A high count could indicate a misunderstanding of the game objective.
- `ShowTime`: number of times a player clicks to display the timer. Regular checks might reflect an active concern for time management as a shared responsibility. `ShowMenu`, `ShowMainGoal` and `ShowTime` could be related to the collective orientation towards game objectives, time management, and strategy planning, all important in co-regulation.
- `RepairsDone`: number of times the player is responsible for repairing a breach. This action depends on the action of a second player, who must be in the same room as the player to help repair the breach.
- `RepairsParticipationDone`: number of times the player has helped repair a breach. This action is linked to the action of the player who initiated the repair of the breach. Both players must be in the same room, and come to the aid of this player so that the repair can be carried out. `RepairsDone` and `RepairsParticipationDone` are strong indicators of collaborative efforts and mutual assistance, key elements in co-regulation.

## 5 Study

### 5.1 Experimental conditions

We compared three conditions that offer different communication modalities for participants who were in separate rooms:

Table 1: Indicators classified by category: task, self-regulation and co-regulation

<b>Task</b>	
<b>Indicator</b>	<b>Description</b>
specialPowerUses	Number of times a special power is used
TurnWithoutScooping	Number of turn without scooping
<b>Self-regulation</b>	
<b>Indicator</b>	<b>Description</b>
ownObjectTaken	Number of times a personal object is retrieved
objectsTaken	Number of times other players' objects are stolen
flashLightUseSum (s)	Total time the flashlight is used
showSubGoal	Number of times player clicks to display sub-goal
<b>Co-regulation</b>	
<b>Indicator</b>	<b>Description</b>
canceledAction	Number of actions canceled
impossibleAction	Number of actions deemed impossible
turnsPlayed	Number of turns played
commonObjectTaken	Number of times the common object is retrieved
meanTurn (s)	Average duration of a turn
showMenu	Number of times player clicks to display menu
showMainGoal	Number of times player clicks to display main-goal
showTime	Number of times player clicks to display timer
repairsDone	Number of times player is responsible for repairing a breach
RepairsParticipationDone	Number of times player has helped repair a breach

- Chat condition: participants were able to communicate via chat.
- Audio condition: participants were able to communicate via audio and chat.
- Video condition: participants were able to communicate via video and chat.

We relied on previous research to select the most commonly studied communication modalities. In their systematic review of the literature, [56] underscore that studies usually compare text-based communication with more vivid and richer forms of communication, in particular audio-only and audiovisual modalities. We decided to maintain text chat available under all conditions to reflect common configurations in contemporary collaborative platforms (e.g. Discord, Zoom, Microsoft Teams) and collaborative games.

## 5.2 Participants

A sample of 72 participants was recruited for this study through the mailing lists of our university. The average age was 21.5 years ( $SD = 1.7$ ) and there was a distribution of 25 men and 47 women.

## 5.3 Procedure

We opted for between-subjects experimental design where participants were assigned to one experimental condition for the all experiment. Upon arrival, the

participants were taken to a designated room where the experiment was introduced. After filling out pre-test questionnaires, they were directed to a tutorial session of the game.

Then, they played two consecutive parts of the game with the same two other participants, guarantying to collect enough long game sessions for the analysis of the evolution of regulation processes (RQ3).

The game content remained identical across all conditions and parts, ensuring comparability. After completing both parts, participants filled in post-test questionnaires.

#### 5.4 Data collection

Our research adopted a multimodal approach to data collection. We used a logging system to capture interaction data throughout the gaming sessions. Each participant was assigned a unique identifier that linked their questionnaire responses, game interactions, and game role. We also captured chat logs, video and audio streams and recorded them using the anonymized identifier, while removing all personally identifying information.

Following the game, we used the Networkeds Mind Social Presence Inventory (SPI, 34 items) [11] to explore participants' perceptions of social presence within the game. This questionnaire, with its focus on co-presence and psychobehavioral accessibility, gauged both "self-perception" and "perception of others". This allowed us to ascertain perceived subjective symmetry among participants.

## 6 Results

### 6.1 Impact of communication modalities on social presence

The descriptive statistics on the results of the social presence questionnaire, for different communication modalities are shown in table 2. As the data deviated from a normal distribution and included three independent samples, non-parametric tests were conducted. The Kruskal-Wallis test was first applied, followed by a Dunn's post hoc test on the social presence questionnaire results across different conditions, as illustrated in tables 4, 5. In terms of perceived subjective symmetry, Spearman's correlation test was conducted for each questionnaire, comparing "self" and "others" dimensions. We calculated the mean coefficients for the three conditions.

*Co-presence:* The Kruskal-Wallis Test on the co-presence dimension of the Networked Minds Social Presence Inventory (NM-SPI) yielded a p-value ( $\rho$ ) of .008, which is below the alpha ( $\alpha$ ) threshold of 0.05, indicating statistically significant results (Table 4). Further analysis through Dunn's Post Hoc comparisons (Table 5) revealed that a significant difference exists between chat and audio conditions, as the p-value ( $\rho = .002$ ) is less than the alpha ( $\alpha = 0.05$ ) threshold. The data suggests that the audio modality facilitates a higher degree of co-presence (M = 5.88; SD = 0.58) compared to the chat modality (M = 5.15;

Table 2: Descriptive statistics regarding the 3 communication modalities according to the dimension of social presence, on a likert scale (1 to 7).

	Co-presence		Psycho-behavioral			
	Audio	Chat	Video	Audio	Chat	Video
Number of participants	23	24	24	23	24	24
Median	5.88	5.06	5.44	5.92	5.63	5.60
Mean	5.88	5.15	5.52	5.87	5.45	5.52
Std. Deviation	0.58	0.88	0.69	0.57	0.63	0.45
Coefficient of variation	0.10	0.17	0.13	0.10	0.12	0.08
Variance	0.34	0.77	0.48	0.33	0.40	0.20

Table 3: Spearman's Correlation Coefficients ( $\rho$ ) for the perceived subjective symmetry dimension

Dimension	Chat		
	Audio	Video	Video
Co-Presence	0.83	0.76	0.72
Psychobehavioral Accessibility	0.82	0.80	0.76

Table 4: Kruskal-Wallis Test for Co-Presence and Psychobehavioral Accessibility dimensions

Dimension	Factor	Statistic	df	p
Co-Presence	Condition	9.633	2	0.008**
Psychobehavioral Accessibility	Condition	7.647	2	0.022*

\*\*  $p < .01$ , \*  $p < .05$

Table 5: Dunn's Post Hoc Comparisons for the Co-presence and Psychobehavioral Accessibility dimensions

Dimension	Comparison	z	p	$p_{bonf}$
Co-presence	Audio - Chat	3.085	0.002**	0.006**
Co-presence	Audio - Video	1.868	0.062	0.185
Co-presence	Chat - Video	-1.230	0.219	0.657
Psychobehavioral	Audio - Chat	2.410	0.016*	0.048*
Psychobehavioral	Audio - Video	2.396	0.017*	0.050*
Psychobehavioral	Chat - Video	-0.014	0.989	1.000

\*\*  $p < .01$ , \*  $p < .05$

SD = 0.88). However, no significant differences were observed between other pairs (audio and video, chat and video). This suggests that participants using the audio condition are more aware of the co-presence of other's mediated body than in chat conditions.

*Psychobehavioral Accessibility.* The Kruskal-Wallis Test produced a p-value ( $\rho$ ) of .022, which is below the alpha ( $\alpha$ ) threshold of 0.05, indicating significant results (Table 4). Further investigation with the Dunn's Post Hoc comparisons (Table 5) showed significant differences between the audio and chat modalities ( $\rho = .016$ ), and the audio and video modalities ( $\rho = .017$ ), both falling below the alpha ( $\alpha = 0.05$ ) threshold. These results suggest that the audio modality enables higher psychobehavioral accessibility ( $M = 5.87$ ;  $SD = 0.57$ ) than both the chat modality ( $M = 5.45$ ;  $SD = 0.63$ ) and the video modality ( $M = 5.52$ ;  $SD = 0.45$ ). However, no significant differences were detected in the other comparisons (chat and video). This suggests that participants using the audio condition have a greater sense of access to the attentional engagement, emotional state, understanding and behavioral interaction of others than in the other two conditions.

*Perceived subjective symmetry.* As illustrated in (Table 3), the mean of the correlation coefficients is larger in the chat modality for both the co-presence dimension (chat:  $\rho = 0.83$ ; audio:  $\rho = 0.76$ , video:  $\rho = 0.72$ ) and the perceived psychobehavioral accessibility dimension (chat:  $\rho = 0.82$ ; audio:  $\rho = 0.80$ , video:  $\rho = 0.76$ ). This suggests that participants using the chat modality demonstrated a better understanding of other participants' perceptions of them compared to those in the third-person condition.

## 6.2 Impact of communication modalities on regulation

**Difference between chat, audio and video conditions** The objective of the analysis was to explore indicators related to regulation processes. This was achieved by performing an ANOVA test on the indicators of the three different conditions that were presented. The selected indicators are presented in Table 6. For our statistical analysis, we normalized all indicators by dividing each participant's actions by their number of game turns, producing a ratio that accounts for varying gameplay durations. These normalized ratios were then averaged per condition.

From the total number of indicators investigated, five of them were identified as significant. They include the number of turns played ( $p < 0.001$ ), the mean turn ( $p < 0.001$ ), impossible actions ( $p < 0.001$ ), actions of showing the menu ( $p < 0.05$ ), and showing the sub-goal ( $p < 0.01$ ). All the other indicators were not significant.

For the significant indicators, further analysis was carried out using independent samples t-tests, the results of which are shown in the table 7. The indicators are classified under two categories of regulation processes: self-regulation and co-regulation. To control for multiple comparisons, p-values were adjusted using the Holm-Bonferroni sequential correction method.

In the self-regulation category, the 'showSubGoal' indicator revealed significant differences between the Audio ( $M = 0.659$ ) and Chat ( $M = 0.289$ ) conditions ( $p < 0.001$ ), and between the Chat ( $M = 0.289$ ) and Video ( $M = 0.665$ ) conditions ( $p < 0.001$ ). No significant difference was observed between the Audio and Video conditions.



Table 6: Results of an Anova performed on indicators to compare regulation processes between the 3 conditions.

Task				
Indicator	Audio Mean	Chat Mean	Video Mean	p-value
specialPowerUses	0.692	0.634	0.707	0.542
TurnWithOutScooping	0.270	0.266	0.258	0.932
self-regulation				
Indicator	Audio Mean	Chat Mean	Video Mean	p-value
ownObjectTaken	0.055	0.049	0.046	0.863
objectsTaken	0.000	0.006	0.011	0.211
flashLightUseSum	20.97	22.02	25.59	0.404
showSubGoal	0.659	0.289	0.665	0.00158**
co-regulation				
Indicator	Audio Mean	Chat Mean	Video Mean	p-value
canceledAction	0.612	0.699	0.732	0.520
impossibleAction	0.314	0.572	0.269	<0.001***
turnsPlayed	5.69	6.69	5.33	<0.001***
commonObjectTaken	0.021	0.013	0.019	0.748
meanTurn	74.57	97.19	93.78	<0.001***
showMenu	1.03	0.815	1.34	0.00473**
showMainGoal	0.312	0.230	0.319	0.403
showTime	1.53	0.900	1.23	0.093

\*\*\* p < .001, \*\* p < .01, \* p < .05

The co-regulation category displayed more significant results across the indicators. For 'turnsPlayed', significant differences were found between the Audio (M = 5.687) and Chat (M = 6.688) conditions (p < 0.001), and between the Chat (M = 6.688) and Video (M = 5.333) conditions (p < 0.001). 'showMenu' revealed a significant difference between Chat (M = 0.815) and Video (M = 1.343) conditions (p < 0.001), whereas 'meanTurn' exhibited significant differences between Audio (M = 74.568) and Chat (M = 97.188) conditions (p < 0.001) and Audio (M = 74.568) and Video (M = 93.781) conditions (p < 0.001). Lastly, 'impossibleAction' indicator had significant differences between Audio (M = 0.314) and Chat (M = 0.572) conditions (p < 0.01) and between Chat (M = 0.572) and Video (M = 0.269) conditions (p < 0.001).

However, in stark contrast to the preceding observations, all the indicators associated with the task category displayed non-significant results. Regardless of the conditions under consideration, the associated p-values were uniformly greater than the accepted threshold, thus failing to reject the null hypothesis.

**Evolution between part 1 and 2** The analysis aimed to explore the evolution of the indicators between part 1 and part 2. To this end, a pairwise t-test was carried out on the indicators and the significant results are presented in Table 8. P-values were adjusted using the Benjamini-Hochberg procedure [7] to control the false discovery rate (FDR). When analyzing the evolution between the two

Table 7: T-test performed on indicators with a significant difference between the 3 conditions

Self-regulation						
Indicator	Condition	$\bar{x}_{Audio}$	$\bar{x}_{Chat}$	$\bar{x}_{Video}$	p-value	adj. p-value
showSubGoal	Audio vs Chat	0.659	0.289		<0.001***	0.00144**
showSubGoal	Audio vs Video	0.659		0.665	0.967	0.967
showSubGoal	Chat vs Video		0.289	0.665	<0.001***	0.00144**
Co-regulation						
Indicator	Condition	$\bar{x}_{Audio}$	$\bar{x}_{Chat}$	$\bar{x}_{Video}$	p-value	adj. p-value
impossibleAction	Audio vs Chat	0.314	0.572		0.002**	0.018*
impossibleAction	Audio vs Video	0.314		0.269	0.513	1.000
impossibleAction	Chat vs Video		0.572	0.269	<0.001***	0.002**
turnsPlayed	Audio vs Chat	5.687	6.688		<0.001***	0.00144**
turnsPlayed	Audio vs Video	5.687		5.333	0.136	0.136
turnsPlayed	Chat vs Video		6.688	5.333	<0.001***	<0.001***
meanTurn	Audio vs Chat	74.568	97.188		<0.001***	<0.001***
meanTurn	Audio vs Video	74.568		93.781	<0.001***	<0.001***
meanTurn	Chat vs Video		97.188	93.781	0.557	0.557
showMenu	Audio vs Chat	1.032	0.815		0.125	0.852
showMenu	Audio vs Video	1.032		1.343	0.089	0.266
showMenu	Chat vs Video		0.815	1.343	<0.001***	0.00268**

\*\*\* p < .001, \*\* p < .01, \* p < .05

parts, several indicators showed differences that, while significant before FDR correction, did not maintain significance after correction. These indicators are presented as observable trends that suggest potential patterns of evolution, while acknowledging their reduced statistical strength.

In the task category, we observed a trend toward increased use of special powers in the chat condition (part 1:  $M = 0.58$ ; part 2:  $M = 0.69$ ), though this did not maintain significance after FDR correction.

Within the self-regulation category, the most notable change was a significant decrease in the 'showSubGoal' indicator for the audio condition (from  $M = 0.93$  to  $M = 0.39$ ,  $p = 0.0003$ ). A tendency to decrease in flashlight usage was also observed in the chat condition.

The co-regulation category showed several robust changes, particularly in turn dynamics. The 'meanTurn' indicator decreased significantly in both audio (from  $M = 84,959.60$  to  $M = 64,176.30$ ,  $p = 0.0001$ ) and chat conditions (from  $M = 106,993.99$  to  $M = 87,382.62$ ,  $p = 0.002$ ). We also observed a trend towards a decrease in the number of turns played in the video condition and in the number of impossible actions in the audio condition.

Interface interaction patterns processes also evolved between parts. We observed a decrease in the audio condition of the 'showMainGoal' indicator (from  $M = 0.47$  to  $M = 0.15$ ,  $p = 0.001$ ), as well as in the 'showTime' indicator in the same condition (from  $M = 2.05$  to  $M = 1.01$ ,  $p = 0.005$ ). Overall, this decreasing

Table 8: Significant differences between part 1 and part 2

Task					
Indicator	Condition	$\bar{x}_1$	$\bar{x}_2$	P-value	adj. p-value
specialPowerUses	chat	0.58	0.69	0.02*	0.129
Self-regulation					
Indicator	Condition	$\bar{x}_1$	$\bar{x}_2$	P-value	adj. p-value
flashLightUseSum	chat	25.415	18.632	0.02*	0.129
showSubGoal	audio	0.93	0.39	0.0003***	0.006**
Co-regulation					
Indicator	Condition	$\bar{x}_1$	$\bar{x}_2$	P-value	adj. p-value
impossibleAction	audio	0.42	0.20	0.0098**	0.075
turnsPlayed	video	5.83	4.83	0.02*	0.129
meanTurn	audio	84.96	64.18	0.0001***	0.006**
meanTurn	chat	106.99	87.38	0.002**	0.028*
showMenu	chat	0.93	0.70	0.01*	0.075
showMainGoal	audio	0.47	0.15	0.001**	0.020*
showMainGoal	video	0.40	0.24	0.04*	0.212
showMainGoal	chat	0.33	0.13	0.02*	0.129
showTime	audio	2.05	1.01	0.005**	0.046*
showTime	chat	1.18	0.62	0.04*	0.212

\*\*\* p < .001, \*\* p < .01, \* p < .05

trend was identified in all conditions for the 'showMenu', 'showMainGoal', and 'showTime' indicators.

These patterns suggest an overall trend toward more efficient interaction across conditions, with particularly robust improvements in the audio condition, though not all trends reached statistical significance after correction for multiple comparisons.

### 6.3 Correlation between social presence and regulation

A series of Spearman correlation tests were conducted to analyze the relationship between indicators and the results obtained from the two dimensions (co-presence and psycho-behavioral) of the Networked Mind Social Presence Inventory. Spearman's correlation was chosen as the method of analysis due to the data not following a normal distribution.

Regarding the "Co-presence" dimension, a positive correlation was observed between 'repairParticipationsDone' for the video condition ( $r = 0.446512$ ,  $p = .028724$ ), 'turnsPlayed' for the chat condition ( $r = 0.436734$ ,  $p = .032858$ ), and 'commonObjectTaken' for the chat condition ( $r = 0.414321$ ,  $p = .044126$ ). These results suggest a medium correlation which are statistically significant at the 0.05 threshold.

However, negative correlations were found between 'showMenu' for the audio condition ( $r = -0.422825$ ,  $p = .039540$ ) and 'canceledAction' for the audio condi-

Table 9: Significant results between indicators and presence dimensions after spearman correlation.

<b>Co-presence</b>			
<b>Indicator</b>	<b>Condition</b>	<b>Correlation</b>	<b>P-value</b>
<i>co-regulation</i>			
canceledAction	audio	-0.500	0.0274*
repairParticipationsDone	video	0.447	0.0287*
turnsPlayed	chat	0.437	0.0329*
commonObjectTaken	chat	0.414	0.0441*
showMenu	audio	-0.423	0.0395*
<b>Psycho-behavioral</b>			
<b>Indicator</b>	<b>Condition</b>	<b>Correlation</b>	<b>P-value</b>
<i>Task</i>			
turnsWithoutScooping	audio	-0.528	0.00800**
<i>self-regulation</i>			
showSubGoal	chat	0.407	0.0481*
<i>co-regulation</i>			
commonObjectTaken	chat	0.505	0.0119*
showMenu	video	0.437	0.0325*
repairsDone	video	0.406	0.0489*

\*\* p < .01, \* p < .05

tion ( $r = -0.449916$ ,  $p = .027386$ ). These results indicate a relationship between a decrease in these indicators and an increase in the co-presence dimension.

For the "Psycho-behavioral" category, all correlations were positive, except for 'turnsWithoutScooping' in the audio condition, which showed a significant negative correlation ( $r = -0.527994$ ,  $p = .008003$ ). This result indicates a relationship between a decrease in this indicator and an increase in the psycho-behavioral dimension. Other positive correlations included 'commonObjectTaken' for the chat condition ( $r = 0.504602$ ,  $p = .011915$ ), 'showMenu' for the video condition ( $r = 0.437407$ ,  $p = .032559$ ), 'showSubGoal' for the chat condition ( $r = 0.407474$ ,  $p = .048113$ ), and 'repairsDone' for the video condition ( $r = 0.406139$ ,  $p = .048922$ ). All of these results indicate medium correlations and are statistically significant at the 0.05 threshold.

## 7 Discussion

This discussion analyzes the varying effects of different communication modalities (audio, chat, and video) used in a collaborative gaming setting. Specifically, we examine how each modality affects participants' self-regulation, coregulation, and sense of social presence.

## 7.1 Audio condition

Our results highlight that the audio condition notably enhanced both the co-presence (i.e. the perception of the other’s mediated body) and psychobehavioral accessibility (i.e. the sense of the accessibility of the other psychological, emotional and intentional states). This is accompanied by more effective co-regulation, notably through indicators such as a lower number of game turns and shorter turn times, as well as fewer impossible actions.

The higher sense of co-presence compared to the chat condition could be explained by the fact that the voice is perceived by participants as a representation of the other’s mediated body [65]. Furthermore, participants in the audio condition experienced a higher sense of psychobehavioral accessibility in comparison to both chat and video conditions, indicating a greater understanding of the emotional and intentional states of other players. This could be explained by the emotional nuance conveyed by tones of voice (not allowed by text communication), as well as a greater facility in expressing to others the access to their own mental states, enabling an improved sense of psycho-behavioral accessibility. These results are in line with studies showing that social presence is higher when participants communicate via audio (explanations on the difference between audio and video are provided in the next section) [8, 2, 44].

From a self-regulation perspective, participants in the audio condition displayed more personal goals than in the chat condition. This suggests that the audio modality may foster an environment where players are more likely to pursue individual strategies and objectives. This may be due to a lower cognitive load than in the chat condition, as shown in the work of Erlandson et al. [22], which allows players to explore secondary goals.

Analyzing the co-regulation processes revealed that the audio condition significantly improved gameplay strategy and efficiency, as evidenced by a decrease in the number of turns, a reduced average time per turn, in comparison to chat and video conditions. The absence of visual cues in the audio condition might have nudged participants to articulate their strategies more clearly, fostering quicker consensus and decision-making. Given the complexity of the game, the audio condition certainly made explanations easier compared to the chat condition [14]. This co-regulation dynamic improved between the first and the second part of the game, indicating a growing familiarity and shared understanding amongst team members.

After analyzing the sense of presence and regulation processes, correlation analysis further elucidated the relationships between collaboration dynamics and the perceived sense of presence. A noticeable negative correlation was observed between the number of menu displays and cancelled actions, and the feeling of co-presence. This suggests that a stronger sense of co-presence might foster better communication, thus reducing occurrences of misunderstandings or poor planning, which are the primary causes of action cancellations. Additionally, a higher degree of psychobehavioral accessibility was correlated with a decreased number of rounds played without scooping, implying that a deeper mutual un-

derstanding [10] might lead to more strategic and efficient gameplay decisions, consequently reducing water accumulation and preventing game losses.

In conclusion, the audio condition seems to foster a higher sense of social presence, characterized by enhanced co-presence and psychobehavioral accessibility. This, in turn, facilitates better self and co-regulation processes, which positively impact the overall game strategy and success rates.

## 7.2 Video condition

Regarding the video condition, we found that social presence is lower than in the audio condition (only for the psychobehavioral dimension). This is associated with certain aspects of co-regulation, which are more effective in the "video" condition than in the "chat" condition, but less effective than in the "audio" condition.

Results indicate a lower sense of social presence, specifically psycho-behavioral accessibility, in the video condition compared to the audio condition. This may seem unexpected at first given that video communication provides more visual cues (e.g. facial expressions, gaze) and non-verbal information that could potentially enhance the sense of accessibility of the other being's psychological, emotional, and intentional states. However, there is a division in the literature on this subject. Some studies suggest that the video modality enhances social presence [19, 18], while others find no improvement [4, 8, 32, 62].

We think that this finding could be explained by the cognitive load theory [57]. The integration of both visual and auditory stimuli in video communication may lead to a higher cognitive load compared to audio communication, which relies only on auditory stimuli [64]. This increased cognitive demand could lead to a lower sense of psycho-behavioral accessibility because the players might find it difficult to simultaneously manage their game actions, visually process non-verbal cues, and maintain a high sense of social presence. It is also possible that the size of the screen allocated to the other players' webcams, as well as their location (outside the game screen, see figure 1 and 2) may have had an impact on this result. It is also possible, as shown in certain studies [19, 18], that the impact of video modality on social presence could be moderated by the nature of the task, e.g. a visual task.

Our results also show a significant difference in co-regulation with a lower number of turns and fewer impossible actions (in comparison with the chat modality) and a longer average playing time per turn (in comparison with the chat and audio modality). This could be an indication of increased efficiency and strategic planning in the video condition, possibly as a consequence of the additional visual cues available, thus potentially reducing the number of necessary turns but lengthening the average time per turn with the additional visual cues adding a higher cognitive load, or a source of distraction [13]. Another explanation is that the video format facilitated additional conversational content, such as expressing agreement, providing encouragement, having fun, and seeking definitions of discussed concepts [61]. There is a significant difference in task-

relatedness, and in self-regulation, there is more display of secondary objectives than in the chat condition.

Upon examining the dynamics of gameplay and their regulatory processes, the correlation analysis further clarified the associations between these dynamics and the perceived sense of social presence. We observe that co-presence is positively correlated with the number of repair aids performed. Repair assists imply that both people are in the same room, and that the person who is playing sends a repair request to the second player, who must confirm it. Although there is no representation of other players in the first-person viewpoint, it seems that performing a repair action involving two people in the same room has an impact on co-presence, i.e. the mediated representation of the other's body. This result applies only to the video condition. Psycho-behavioral accessibility is positively correlated with the number of repairs made by the player. It is possible that taking joint action with another player, by initiating the request, improves the perception of access to the other player. As this correlation was only found in the video condition, it is possible that it is related to viewing other players via the webcam. This phenomenon can be explained by the fact that, in this condition, players are visible via a webcam but have no representation when they are in the same room in the game. The addition of common actions and highlighting the presence of the other player in the same room may change their perception of the other players linked to the webcam. Biocca and Harms [10] argue that there is a threshold social presence moment "when the form of a medium (light reflecting inks, pixels, or marble) moves from being a thing, matter, to being social, a representation of another."

In conclusion, the video condition appears to lead to a lower sense of social presence than the audio condition, characterized by lower psychobehavioral accessibility. Some elements of co-regulation are better than the chat condition (i.e. lower number of turns, fewer impossible actions), but only a few.

### 7.3 Chat condition

In our study, we found that in the chat condition, social presence is lower than in the audio condition for both dimensions. This is accompanied by less effective co-regulation, compared with the audio and video conditions. There is no significant difference in the task-related indicators, and in the self-regulation indicators, there is less display of sub-goals.

We found that the chat condition yielded a significantly lower sense of social presence, specifically with regards to both co-presence and psycho-behavioral accessibility, in comparison to the audio conditions. This suggests that text communication may be less effective in fostering the feeling of "being with others", potentially due to the lack of non-verbal cues or real-time emotional expressions often available in audio and video modalities. This finding is consistent with prior literature indicating that text communication often results in a reduced sense of social presence due to its inherent limitations (e.g., time delays in response, absence of tonal nuances, and absence of visible emotional cues) [56], as well as with our previous explanations regarding the cognitive load.

Analyzing the co-regulation processes revealed that the chat condition also exhibited significantly more turns, indicating that the development of a cooperative strategy might take longer in a chat environment compared to audio or video. This could be attributed to the asynchronous nature of chat communication, which may limit the efficiency of real-time interaction and feedback, crucial for effective co-regulation [39]. Simultaneously, we observed that players in the chat condition had a significantly longer average playing time per turn compared to those in the audio condition. These results are in line with the study of Sallnäs [62] that shows that the time spent on a decision-making task was longer in a text chat environment, compared to an audio or video condition. However, the average turn duration significantly decreased between the first and second part of the game, suggesting that players may have begun to adapt their co-regulation strategies to the chat-based environment, thereby enhancing their co-regulation processes efficiency over time. Moreover, participants under the chat condition performed significantly more impossible actions compared to their counterparts in the audio and video conditions, indicating poorer co-regulation. This might be due to the limited contextual and emotional understanding afforded by the text modality, which could impede effective joint problem-solving and decision-making, making it more difficult to provide complex explanations [14].

In the realm of self-regulation, players in the chat condition displayed their personal goals less frequently than those in the audio and video conditions. The lack of visual and audio cues in chat communication could potentially reduce the motivation or perceived need to share personal goals. This could have the impact of further undermining self-regulation behaviors. It is also possible that the cognitive load of the chat condition limits players' interest in secondary objectives due to a lack of time or attention span.

Our correlation analysis provided additional insights into the links between these gameplay elements and the player's perceived sense of presence. Interestingly, we found a significant positive correlation between the number of turns played and the sense of co-presence within the chat condition. This might suggest that increased number of turns in the game could help compensate for the initial lack of social presence in chat environments. Thus, players may progressively develop a sense of co-presence as they adapt to the text communication modality, as shown in the work of Ramirez et al. [60].

Moreover, a significant positive correlation was observed between the instances of common object retrieval and both the sense of co-presence and psycho-behavioral accessibility. This result implies that shared activities or goals might enhance the perceived social presence and accessibility in chat-based interactions. This finding aligns with prior research suggesting that joint activities could serve as a mechanism for fostering a sense of togetherness and shared understanding in collaborative tasks [36].

These results are also accompanied by a significant evolution in co-regulation between the first and second parts, notably with shorter turnaround times, and fewer displays of main goals and time.



In conclusion, the chat condition appears to lead to a lower sense of social presence than the audio condition, characterized by lower co-presence and psychobehavioral accessibility. However, the results show an improvement in certain indicators between the first and second games, suggesting that the players have adapted to this modality. The various indicators are also weaker than in other conditions.

#### 7.4 Implications for design

Given the insights of this study, we derive some recommendations for the design of collaborative environments. First, we believe that designers should prioritize integrating robust audio communication channels to bolster the sense of co-presence and psychobehavioral accessibility among participants. Furthermore, developers should consider strategies to minimize the cognitive load associated with text-based interactions to allow players more scope to explore secondary goals and strategies, enhancing their collaboration. In collaborative environments, these findings suggest a potential avenue for improving remote collaboration and learning through the use of audio channels that enhance psychobehavioral accessibility. This can foster a better understanding and appreciation of individual perspectives and emotional nuances, thus enriching the collaborative experience.

To augment the observed benefits of audio conditions in fostering co-presence and improving gameplay strategy, it would be beneficial to integrate mechanisms that encourage participants to express their personal goals vocally. This could be achieved through the design of game mechanics that reward clear verbal strategizing and goal setting, which could potentially lead to a richer and more collaborative gaming experience.

Additionally, efforts should be directed towards optimizing the gaming interface to reduce actions that negatively correlate with the sense of co-presence, for example to reduce the number of menu displays and action cancellations, thereby facilitating smoother gameplay and improved communication. Moreover, leveraging on the enhanced psycho-behavioral accessibility noted in the audio condition, game designers should consider incorporating features that promote a deeper mutual understanding and strategic collaboration amongst team members, ultimately leading to better collaboration.

Furthermore, it would be interesting to explore further research to discern the optimal balance between audio and video modalities in fostering both co-presence and efficient co-regulation processes. This could pave the way for the development of more nuanced and adaptable collaborative platforms that leverage the strengths of various communication modalities to foster enriched and successful collaborative experiences.

Finally, when employing the video condition, it is important to carefully consider its placement and the size shown to avoid diminishing the user experience. Regarding the chat condition, while some elements may not be engaging, it could be employed for other purposes, such as adding an extra level of difficulty in strategy games or puzzle games.

### 7.5 Limitations

Our study presents several methodological limitations. First, while our analysis of log-derived quantitative indicators provided objective behavioral measures, a more comprehensive understanding could be achieved by triangulating data sources: behavioral logs, user perception questionnaires, and qualitative analysis of communication content. Second, individual differences (such as social skills, prior experience, or personal preferences) were not accounted for, potentially masking important moderating effects on how participants engage with different communication modalities. Third, our findings are specific to a collaborative decision-making game context, and generalization to other collaborative tasks or environments would require further validation.

## 8 Conclusion and Future Work

Overall, our findings highlight the importance of communication modalities in fostering social presence, specifically co-presence and psycho-behavioral accessibility, as well as co-regulation and self-regulation processes in collaborative tasks within virtual environments. We more particularly revealed distinct patterns across communication modalities. Audio communication demonstrated superior performance in fostering social presence and regulation processes, likely due to its optimal balance between information richness and cognitive load. While video communication showed lower psycho-behavioral accessibility, it successfully supported strategic planning and team cohesion through shared actions. Text-based chat, despite showing the lowest initial social presence, exhibited adaptation effects as participants developed compensatory strategies over time. These results suggest that the effectiveness of communication modalities in collaborative virtual environments depends on their ability to balance social presence support with cognitive demands.

As previously described, our study presents several methodological limitations that outline future research opportunities. Further research should focus on developing comprehensive frameworks that capture the complexity of multimodal communication patterns in collaborative environments. This includes creating and validating tools that measure regulation processes from both objective and subjective perspectives, providing a more complete understanding of collaborative dynamics. Additionally, qualitative analysis of communication content across different modalities would reveal deeper insights into how various channels influence interaction patterns and effectiveness. Future work should also investigate how individual characteristics such as social skills, prior experience, and communication preferences moderate the relationship between communication modalities and collaborative outcomes.

Beyond methodological improvements, our findings point to broader research directions in computer-mediated collaboration. First, further investigation into the nuanced effects of audio modality on group dynamics and individual behavior in virtual environments is needed. Given the observed benefits of audio in enhancing co-presence and psycho-behavioral accessibility, understanding its

influence on various group sizes, task complexities, and cultural contexts could yield valuable insights. Second, exploring the potential of video communication in more detail, particularly in terms of balancing cognitive load and fostering strategic collaboration, would be beneficial. Studies could examine how varying video quality, screen layouts, and the presence of visual cues impact group interaction and task performance. Third, the adaptation observed in the chat condition opens a new avenue for research into how users evolve their communication strategies over time in text-based environments. Longitudinal studies could provide a deeper understanding of the learning curve and the development of compensatory mechanisms for the lack of non-verbal cues. Finally, integrating multimodal communication, combining audio, video, and text, in collaborative virtual settings presents a promising area for research.

## 9 Ethical

The Research Ethics Committee of the University of Lyon was only established recently in January 2022. By this date, we had already developed our protocol and consent forms for participants, and the requirement for approval from this committee was not yet instituted in our practices. Nevertheless, our approach is in compliance with the fundamental principles outlined in the Declaration of Helsinki concerning research involving human beings, as well as the ethical guidelines of the American Psychological Association (APA). We ensured that our research adhered to these principles by providing complete and transparent information to participants, guaranteeing their right to confidentiality and anonymity, and implementing mechanisms to minimize any potential risk associated with their participation. The Data Protection Officer, affiliated with the Ecole Normale Supérieure de Lyon (ENS), supervised the creation and administration of the consent forms, thereby complying with the General Data Protection Regulation (GDPR) and ensuring that all participants were fully informed about the nature of the study and gave their informed consent. Our study did not involve any invasive procedures since we only collected interaction traces and questionnaires. All collected data were fully anonymized to protect the identity and confidentiality of the participants, and they were free to opt-out of data collection if they wished. The participants' interaction with the study was limited to playing a non-violent, collaborative game designed to be played in turns, without any visual effects that could induce discomfort or disorders in individuals with certain sensitivities. The study's conditions were explained in advance, and participants were free to withdraw at any time. Considering these elements, we have acted responsibly and in accordance with current ethical expectations, ensuring the protection of participants and respecting their consent and anonymity. We hope this explanation provides reviewers and readers with a clear understanding of the measures we have taken to guarantee the ethical integrity of our research.

## References

1. Akyol, Z., Garrison, D.: The development of a Community of Inquiry over time in an online course: Understanding the progression and integration of social, cognitive and teaching presence. *Journal of Asynchronous Learning Networks* **12**, 3–22 (2008). <https://doi.org/10.24059/olj.v12i3.72>
2. Appel, J., von der Pütten, A., Krämer, N.C., Gratch, J.: Does Humanity Matter? Analyzing the Importance of Social Cues and Perceived Agency of a Computer System for the Emergence of Social Reactions during Human-Computer Interaction. *Advances in Human-Computer Interaction* **2012** (2012). <https://doi.org/10.1155/2012/324694>
3. Bachour, K., Kaplan, F., Dillenbourg, P.: An Interactive Table for Supporting Participation Balance in Face-to-Face Collaborative Learning. *IEEE Transactions on Learning Technologies* **3**(3), 203–213 (2010). <https://doi.org/10.1109/TLT.2010.18>
4. Bailenson, J.N., Yee, N., Merget, D., Schroeder, R.: The Effect of Behavioral Realism and Form Realism of Real-Time Avatar Faces on Verbal Disclosure, Nonverbal Disclosure, Emotion Recognition, and Copresence in Dyadic Interaction. *Presence: Teleoperators and Virtual Environments* **15**(4), 359–372 (2006). <https://doi.org/10.1162/pres.15.4.359>
5. Baltes, B.B., Dickson, M.W., Sherman, M.P., Bauer, C.C., LaGanke, J.S.: Computer-Mediated Communication and Group Decision Making: A Meta-Analysis. *Organizational Behavior and Human Decision Processes* **87**(1), 156–179 (2002). <https://doi.org/10.1006/obhd.2001.2961>
6. Basille, A., Lavoué, , Serna, A.: Impact of Viewpoint on Social Presence and Collaborative Processes in a Collaborative Serious Game. In: *Proceedings of the 14th International Conference on Computer Supported Education. Proceedings of the 14th International Conference on Computer Supported Education*, vol. 2, pp. 614–621. Prague, Czech Republic (Apr 2022). <https://doi.org/10.5220/0011046400003182>, <https://hal.science/hal-03607687>
7. Benjamini, Y., Hochberg, Y.: Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**(1), 289–300 (Jan 1995). <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>, <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
8. Bente, G., Rüggenberg, S., Krämer, N.C., Eschenburg, F.: Avatar-Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net-Based Collaborations. *Human Communication Research* **34**(2), 287–318 (2008). <https://doi.org/10.1111/j.1468-2958.2008.00322.x>
9. Bergstrom, T., Karahalios, K.: Conversation Clock: Visualizing audio patterns in co-located groups. In: *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. pp. 78–78 (2007). <https://doi.org/10.1109/HICSS.2007.151>
10. Biocca, F., Harms, C.: Defining and measuring social presence: Contribution to the networked minds theory and measure. In: In F.R. Gouveia, & F. Biocca (Eds). *Proceedings of the 5th International Workshop on Presence*. pp. 7–36 (2002)
11. Biocca, F., Harms, C., Burgoon, J.K.: Toward a More Robust Theory and Measure of Social Presence: Review and Suggested Criteria. *Presence: Teleoperators and Virtual Environments* **12**(5), 456–480 (2003). <https://doi.org/10.1162/105474603322761270>

12. Bodemer, D., Dehler, J.: Group awareness in CSCL environments. *Computers in Human Behavior* **27**(3), 1043–1045 (2011). <https://doi.org/10.1016/j.chb.2010.07.014>
13. Brünken, R., Plass, J.L., Leutner, D.: Assessment of Cognitive Load in Multimedia Learning with Dual-Task Methodology: Auditory Load and Modality Effects. *Instructional Science* **32**(1/2), 115–132 (2004), <https://www.jstor.org/stable/41953640>
14. Bullard, C., Kansal, A., Freeman, J.: Comparing Chat Methods for Remote Collaborative Live-Coding Music. In: Proceedings of the 16th International Audio Mostly Conference. pp. 200–203. AM '21, Association for Computing Machinery, New York, NY, USA (2021). <https://doi.org/10.1145/3478384.3478399>
15. Casale, S., Fiovaranti, G., Caplan, S.: Online Disinhibition. *Journal of Media Psychology* **27**(4), 170–177 (2015). <https://doi.org/10.1027/1864-1105/a000136>
16. Chandrasegaran, S., Bryan, C., Shidara, H., Chuang, T.Y., Ma, K.L.: TalkTraces: Real-Time Capture and Visualization of Verbal Content in Meetings. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. pp. 1–14. ACM, Glasgow Scotland Uk (2019). <https://doi.org/10.1145/3290605.3300807>
17. Daft, R.L., Lengel, R.H.: Organizational information requirements, media richness and structural design. *Management Science* **32**(5), 554–571 (1986). <https://doi.org/10.1287/mnsc.32.5.554>
18. De Greef, P.: Video communication best for female friends?: ISPR 2014 : 15th International Workshop on Presence (PRESENCE 2014). Challenging Presence: Proceedings of the International Society for Presence Research (ISPR 2014), 17-19 March 2014, Wien, Austria pp. 187–193 (2014)
19. De Greef, P., Ijsselstein, W.A.: Social Presence in a Home Tele-Application. *CyberPsychology & Behavior* **4**(2), 307–315 (2001). <https://doi.org/10.1089/109493101300117974>
20. Dede, C.: Comparing frameworks for 21st century skills. In: In J. Bellanca & R. Brandt (Eds.). *21st Century Skills: Rethinking How Students Learn*. Bloomington, IN: Solution Tree Press; 2010: 51 – 75
21. Dillenbourg, P.: Collaborative Learning: Cognitive and Computational Approaches. *Advances in Learning and Instruction Series*. Elsevier Science, Inc (1999), eRIC Number: ED437928
22. Erlandson, B.E., Nelson, B.C., Savenye, W.C.: Collaboration modality, cognitive load, and science inquiry learning in virtual inquiry environments. *Educational Technology Research and Development* **58**(6), 693–710 (2010), <https://www.jstor.org/stable/40929473>
23. Fjermestad, J.: An analysis of communication mode in group support systems research. *Decision Support Systems* **37**(2), 239–263 (2004). [https://doi.org/10.1016/S0167-9236\(03\)00021-6](https://doi.org/10.1016/S0167-9236(03)00021-6)
24. Gaver, W.W.: Affordances for interaction: The social is material for design. *Ecological psychology* **8**(2), 111–129 (1996)
25. Goffman, E.: *Behavior in public places: notes on the social organization of gatherings*. Free Press [u.a.], New York, 4. print edn. (1963)
26. Gros, B.: Instructional design for Computer-Supported Collaborative Learning in primary and secondary school. *Computers in Human Behavior* **17**(5), 439–451 (2001). [https://doi.org/10.1016/S0747-5632\(01\)00016-4](https://doi.org/10.1016/S0747-5632(01)00016-4)
27. Hadwin, A., Järvelä, S., Miller, M.: Self-regulation, co-regulation and shared regulation in collaborative learning environments. (2017)

28. Hadwin, A.F., Järvelä, S., Miller, M.: Self-regulated, co-regulated, and socially shared regulation of learning. In: Handbook of self-regulation of learning and performance, pp. 65–84. Educational psychology handbook series, Routledge/Taylor & Francis Group, New York, NY, US (2011)
29. Heeter, C.: Being There: The Subjective Experience of Presence. *Presence: Teleoperators and Virtual Environments* **1**, 262 (1992). <https://doi.org/10.1162/pres.1992.1.2.262>
30. Hertz-Lazarowitz, R., Bar-Natan, I.: Writing development of Arab and Jewish students using cooperative learning (CL) and computer-mediated communication (CMC). *Computers & Education* **39**(1), 19–36 (2002). [https://doi.org/10.1016/S0360-1315\(02\)00019-2](https://doi.org/10.1016/S0360-1315(02)00019-2)
31. Hobman, E.V., Bordia, P., Irmer, B., Chang, A.: The Expression of Conflict in Computer-Mediated and Face-To-Face Groups. *Small Group Research* **33**(4), 439–465 (Aug 2002). <https://doi.org/10.1177/104649640203300403>
32. Homer, B.D., Plass, J.L., Blake, L.: The effects of video on cognitive load and social presence in multimedia-learning. *Computers in Human Behavior* **24**(3), 786–797 (2008). <https://doi.org/10.1016/j.chb.2007.02.009>
33. Huang, X.: Improving Communicative Competence through Synchronous Communication in Computer-Supported Collaborative Learning Environments: A Systematic Review. *Education Sciences* **8**(1), 15 (2018). <https://doi.org/10.3390/educsci8010015>
34. Janssen, J., Erkens, G., Kanselaar, G., Jaspers, J.: Visualization of participation: Does it contribute to successful computer-supported collaborative learning? *Computers & Education* **49**(4), 1037–1065 (2007). <https://doi.org/10.1016/j.compedu.2006.01.004>
35. Johnson, D.W., Johnson, R.T.: Learning Together and Alone: Cooperative, Competitive, and Individualistic Learning. Allyn and Bacon, 5th ed. edn. (1999)
36. Johnson, D.W., Johnson, R.T.: An Educational Psychology Success Story: Social Interdependence Theory and Cooperative Learning. *Educational Researcher* **38**(5), 365–379 (2009). <https://doi.org/10.3102/0013189X09339057>
37. Järvelä, S., Hadwin, A.F.: New Frontiers: Regulating Learning in CSCL. *Educational Psychologist* **48**(1), 25–39 (2013). <https://doi.org/10.1080/00461520.2012.748006>
38. Järvelä, S., Järvenoja, H., Veermans, M.: Understanding the dynamics of motivation in socially shared learning. *International Journal of Educational Research* **47**(2), 122–135 (2008). <https://doi.org/10.1016/j.ijer.2007.11.012>
39. Järvelä, S., Kirschner, P.A., Panadero, E., Malmberg, J., Phielix, C., Jaspers, J., Koivuniemi, M., Järvenoja, H.: Enhancing Socially Shared Regulation in Collaborative Learning Groups: Designing for CSCL Regulation Tools. *Educational Technology Research and Development* **63**(1), 125–142 (2015). <https://www.learntechlib.org/p/152447/>
40. Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., Kirschner, P.A.: What multimodal data can tell us about the students’ regulation of their learning process? *Learning and Instruction* **72**, 101203 (2021). <https://doi.org/10.1016/j.learninstruc.2019.04.004>
41. Järvelä, S., Malmberg, J., Koivuniemi, M.: Recognizing socially shared regulation by using the temporal sequences of online chat and logs in CSCL. *Learning and Instruction* **42**, 1–11 (2016). <https://doi.org/10.1016/j.learninstruc.2015.10.006>
42. Järvelä, S., Volet, S., Järvenoja, H.: Research on Motivation in Collaborative Learning: Moving Beyond the Cognitive–Situative Divide and Combin-

- ing Individual and Social Processes. *Educational Psychologist* **45**(1) (2010). <https://doi.org/10.1080/00461520903433539>
43. Kiesler, S., Siegel, J., McGuire, T.W.: Social psychological aspects of computer-mediated communication. *American Psychologist* **39**(10), 1123–1134 (1984). <https://doi.org/10.1037/0003-066X.39.10.1123>
  44. Kim, H., Suh, K.S., Lee, U.K.: Effects of collaborative online shopping on shopping experience through social and relational perspectives. *Information & Management* **50**(4), 169–180 (2013). <https://doi.org/10.1016/j.im.2013.02.003>
  45. Kim, J., Truong, K.P., Charisi, V., Zaga, C., Lohse, M., Heylen, D., Evers, V.: Vocal turn-taking patterns in groups of children performing collaborative tasks: an exploratory study. In: *Interspeech 2015*. pp. 1645–1649. ISCA (2015). <https://doi.org/10.21437/Interspeech.2015-380>
  46. Kim, J., Kim, H., Tay, B.K., Muniyandi, M., Srinivasan, M.A., Jordan, J., Mortensen, J., Oliveira, M., Slater, M.: Transatlantic Touch: A Study of Haptic Collaboration over Long Distance. *Presence: Teleoperators and Virtual Environments* **13**(3), 328–337 (2004). <https://doi.org/10.1162/1054746041422370>
  47. Kim, T., Chang, A., Holland, L., Pentland, A.S.: Meeting mediator: enhancing group collaboration using sociometric feedback. In: *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. pp. 457–466. CSCW '08, Association for Computing Machinery, New York, NY, USA (2008). <https://doi.org/10.1145/1460563.1460636>
  48. Kirschner, P.A., Kirschner, F., Janssen, J.: The Collaboration Principle in Multimedia Learning. In: Mayer, R.E. (ed.) *The Cambridge Handbook of Multimedia Learning*, pp. 547–575. Cambridge Handbooks in Psychology, Cambridge University Press, Cambridge, 2 edn. (2014). <https://doi.org/10.1017/CBO9781139547369.027>
  49. Kivunja, C.: Exploring the Pedagogical Meaning and Implications of the 4Cs "Super Skills" for the 21st Century through Bruner's 5E Lenses of Knowledge Construction to Improve Pedagogies of the New Learning Paradigm (2015). <https://doi.org/10.4236/ce.2015.62021>
  50. Kreijns, K., Kirschner, P.A., Jochems, W.: Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in Human Behavior* **19**(3), 335–353 (2003). [https://doi.org/10.1016/S0747-5632\(02\)00057-2](https://doi.org/10.1016/S0747-5632(02)00057-2)
  51. Lee, K.M.: Presence, Explicated. *Communication Theory* **14**(1), 27–50 (2004). <https://doi.org/10.1111/j.1468-2885.2004.tb00302.x>
  52. Lipponen, L., Rahikainen, M., Lallimo, J., Hakkarainen, K.: Patterns of participation and discourse in elementary students' computer-supported collaborative learning. *Learning and Instruction* **13**(5), 487–509 (2003). [https://doi.org/10.1016/S0959-4752\(02\)00042-7](https://doi.org/10.1016/S0959-4752(02)00042-7)
  53. Mayne, L.A., Wu, Q.: Creating and measuring social presence in online graduate nursing courses. *Nursing Education Perspectives* **32**(2), 110–114 (Apr 2011). <https://doi.org/10.5480/1536-5026-32.2.110>
  54. Molenaar, I., Chiu, M.M.: Dissecting sequences of regulation and cognition: statistical discourse analysis of primary school children's collaborative learning. *Metacognition and Learning* **9**(2), 137–160 (2014). <https://doi.org/10.1007/s11409-013-9105-8>
  55. Nistor, N., Hernández-García, : What types of data are used in learning analytics? An overview of six cases. *Computers in Human Behavior* **89**, 335–338 (Dec 2018). <https://doi.org/10.1016/j.chb.2018.07.038>

56. Oh, C.S., Bailenson, J.N., Welch, G.F.: A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI* **5** (2018), <https://www.frontiersin.org/articles/10.3389/frobt.2018.00114>
57. Plass, J.L., Moreno, R., Brünken, R.: *Cognitive Load Theory*. Cambridge University Press (2010)
58. Praharaj, S., Scheffel, M., Drachsler, H., Specht, M.: Group Coach for Co-located Collaboration. In: Scheffel, M., Broisin, J., Pammer-Schindler, V., Ioannou, A., Schneider, J. (eds.) *Transforming Learning with Meaningful Technologies*. pp. 732–736. *Lecture Notes in Computer Science*, Springer International Publishing, Cham (2019). [https://doi.org/10.1007/978-3-030-29736-7\\_77](https://doi.org/10.1007/978-3-030-29736-7_77)
59. Praharaj, S., Scheffel, M., Schmitz, M., Specht, M., Drachsler, H.: Towards Automatic Collaboration Analytics for Group Speech Data Using Learning Analytics. *Sensors* **21**(9), 3156 (2021). <https://doi.org/10.3390/s21093156>
60. Ramirez, Jr., A., Walther, J.B., Burgoon, J.K., Sunnafrank, M.: Information-Seeking Strategies, Uncertainty, and Computer-Mediated Communication: Toward a Conceptual Model. *Human Communication Research* **28**(2), 213–228 (2002). <https://doi.org/10.1111/j.1468-2958.2002.tb00804.x>
61. Roos, C.A., Postmes, T., Koudenburg, N.: The microdynamics of social regulation: Comparing the navigation of disagreements in text-based online and face-to-face discussions. *Group Processes & Intergroup Relations* **23**(6), 902–917 (2020). <https://doi.org/10.1177/1368430220935989>
62. Sallnäs, E.L.: Effects of Communication Mode on Social Presence, Virtual Presence, and Performance in Collaborative Virtual Environments. *Presence: Teleoperators and Virtual Environments* **14**(4), 434–449 (2005). <https://doi.org/10.1162/105474605774785253>
63. Schunk, D.H., Zimmerman, B.J.: *Motivation and self-regulated learning: Theory, research, and applications*. Motivation and self-regulated learning: Theory, research, and applications, Lawrence Erlbaum Associates Publishers, Mahwah, NJ, US (2008), pages: xi, 416
64. Sekiyama, K., Sugita, Y.: Auditory-visual speech perception examined by functional MRI and reaction time. *The Journal of the Acoustical Society of America* **110**(5\_Supplement), 2706 (2001). <https://doi.org/10.1121/1.4777337>
65. Sidtis, D., Kreiman, J.: In the Beginning Was the Familiar Voice: Personally Familiar Voices in the Evolutionary and Contemporary Biology of Communication. *Integrative Psychological and Behavioral Science* **46**(2), 146–159 (2012). <https://doi.org/10.1007/s12124-011-9177-4>
66. Skalski, P., Tamborini, R.: The Role of Social Presence in Interactive Agent-Based Persuasion. *Media Psychology* **10**(3), 385–413 (2007). <https://doi.org/10.1080/15213260701533102>
67. Skalski, P., Whitbred, R.: Image versus sound: A comparison of formal feature effects on presence and video game enjoyment. *PsychNology Journal* **8**(1), 67–84 (2010), place: Italy Publisher: PsychNology Journal
68. Stahl, G.: Building Collaborative Knowing : elements of a social theory of CSCL. In: Stribos, J.W., Kirschner, P.A., Martens, R.L. (eds.) *What We Know About CSCL: And Implementing It In Higher Education*, pp. 53–85. *Computer-Supported Collaborative Learning Series*, Springer Netherlands, Dordrecht (2004). [https://doi.org/10.1007/1-4020-7921-4\\_3](https://doi.org/10.1007/1-4020-7921-4_3)
69. Straus, S.G.: Technology, Group Process, and Group Outcomes: Testing the Connections in Computer-Mediated and Face-to-Face Groups. *Human-Computer Interaction* **12**(3), 227–266 (1997). [https://doi.org/10.1207/s15327051hci1203\\_1](https://doi.org/10.1207/s15327051hci1203_1)



70. Straus, S.G., McGrath, J.E.: Does the medium matter? The interaction of task type and technology on group performance and member reactions. *Journal of Applied Psychology* **79**(1), 87–97 (1994). <https://doi.org/10.1037/0021-9010.79.1.87>
71. Sun, C., Shute, V.J., Stewart, A., Yonehiro, J., Duran, N., D’Mello, S.: Towards a generalized competency model of collaborative problem solving. *Computers Education* **143**, 103672 (2020). <https://doi.org/10.1016/j.compedu.2019.103672>
72. Swan, K., Garrison, D.R., Richardson, J.C., Swan, K., Garrison, D.R., Richardson, J.C.: *A Constructivist Approach to Online Learning: The Community of Inquiry Framework* (2009), <https://www.igi-global.com/gateway/chapter/23488>
73. Thompson, L.F., Coovert, M.D.: Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction and postdiscussion accuracy. *Group Dynamics: Theory, Research, and Practice* **7**(2), 135–151 (2003). <https://doi.org/10.1037/1089-2699.7.2.135>
74. Tidwell, L.C., Walther, J.B.: Computer-Mediated Communication Effects on Disclosure, Impressions, and Interpersonal Evaluations: Getting to Know One Another a Bit at a Time. *Human Communication Research* **28**(3), 317–348 (2002). <https://doi.org/10.1111/j.1468-2958.2002.tb00811.x>
75. Vauras, M., Iiskala, T., Kajamies, A., Kinnunen, R., Lehtinen, E.: Shared-regulation and motivation of collaborating peers: A case analysis. *Psychologia: An International Journal of Psychology in the Orient* **46**(1), 19–37 (2003). <https://doi.org/10.2117/psysoc.2003.19>
76. Viberg, O., Hatakka, M., Bälter, O., Mavroudi, A.: The current landscape of learning analytics in higher education. *Computers in Human Behavior* **89**, 98–110 (2018). <https://doi.org/10.1016/j.chb.2018.07.027>
77. Walther, J.B.: Interpersonal effects in computer-mediated interaction: A relational perspective. *Communication Research* **19**(1), 52–90 (1992). <https://doi.org/10.1177/009365092019001003>
78. Walther, J.B.: Computer-Mediated Communication: Impersonal, Interpersonal, and Hyperpersonal Interaction. *Communication Research* **23**(1), 3–43 (1996). <https://doi.org/10.1177/009365096023001001>
79. Wu, S., Lin, T.C., Shih, J.F.: Examining the antecedents of online disinhibition. *Information Technology & People* **30**(1), 189–209 (2017). <https://doi.org/10.1108/ITP-07-2015-0167>
80. Zhao, H., Sullivan, K.P.H., Mellenius, I.: Participation, interaction and social presence: An exploratory study of collaboration in online peer review groups. *British Journal of Educational Technology* **45**(5), 807–819 (2014), <https://eprints.whiterose.ac.uk/121111/>
81. Zimmerman, B.J.: Self-Regulated Learning and Academic Achievement: An Overview. *Educational Psychologist* **25**(1), 3–17 (1990). <https://doi.org/10.1207/s15326985ep25012>