Meeting the future

Comparing Social Presence in Virtual Reality, Video Conferencing, and Face-to-Face

Master Thesis

Y.A. Hoogendoorn

S1982699

Master of Business Administration – Digital Business & Analytics

University of Twente

CGI Supervisor: Chris Seelig

1st Supervisor: Dr. R. Effing

2nd Supervisor: Dr. G.W.J. Bruinsma

23rd of April 2025

Abstract

The shift toward hybrid and remote work has raised important questions about how meeting modalities influence social presence, engagement, and collaboration. Organizations increasingly rely on platforms like Microsoft Teams (MS Teams); however, these platforms often fail to replicate the dynamics of face-to-face (F2F) meetings. Virtual Reality (VR) has emerged as a promising alternative, offering immersive and interactive features, though its practical effectiveness remains uncertain. Therefore, this study examined how perceived social presence and meeting effectiveness differ across VR, MS Teams, and F2F meetings within a professional context at CGI. In a repeated-measures design, seven participants engaged in the same meeting across all three modalities. Data were collected through an adapted version of the Networked Minds Social Presence Inventory and structured observations. Statistical analyses (Repeated Measures ANOVA and Linear Mixed Models) revealed that F2F meetings consistently scored the highest, followed by MS Teams. VR scored significantly lower, particularly in co-presence and perceived affective understanding. Observational findings supported these results, highlighting passive participation and limited non-verbal communication in VR. These outcomes were linked to technical constraints, such as outdated hardware and the lack of eye or facial tracking. The study concludes that, under the tested conditions, VR is not yet a viable substitute for F2F or video conferencing in professional contexts. Practical implications emphasize the need for improved hardware, meeting design, and user training to enhance the future potential of VR in remote collaboration.

Keywords: social presence, virtual reality, video conferencing, face-to-face meetings

Table of contents

1 Introduction

The concept of virtual collaboration has undergone significant evolution in recent years in response to the increasing number of companies adopting hybrid working models. According to a report by McKinsey & Company (2022), 58% of U.S. employees now have the opportunity to work remotely at least once per week, with 35% working entirely remotely or in a hybrid setting. A similar shift occurs in Europe, where 44% of workers engage in hybrid work (Eurofound, 2022). Organizations are increasingly adopting video conferencing platforms such as Microsoft Teams (MS Teams) and Zoom to support this transition. However, despite its widespread adoption, video conferencing cannot replicate the natural social dynamics and engagement levels found in face-to-face (F2F) interactions, limiting its effectiveness.

One key factor to measure the effectiveness of virtual communication is social presence. Short, Williams, and Christie (1976) originally defined social presence as the extent to which a communication medium enables individuals to perceive each other as "real" in mediated environments. In F2F communication, social presence is naturally high due to nonverbal cues, shared spatial awareness, and synchronous interactions (Biocca et al., 2003). In contrast, video conferencing often fails to fully convey non-verbal cues, emotional engagement, and spatial depth, making virtual communication less immersive. Social presence is a key predictor of engagement, collaboration quality, and communication effectiveness in digital environments, highlighting its importance for evaluating the success of remote meetings (Cummings & Bailenson, 2016).

MS Teams and similar platforms face four main limitations in replicating the experience of F2F meetings. These include reduced access to non-verbal cues, lack of spatial awareness, multitasking susceptibility, and lack of depth perception. These limitations hinder emotional connection, engagement, and collaboration during virtual meetings, which are all closely linked to social presence. As discussed in the theoretical framework (Chapter 3), these limitations provide the foundation for exploring alternative technologies that closely replicate F2F meetings' engagement and dynamic interactions.

One technology that could address these shortcomings is Virtual Reality (VR). As an emerging technology, VR has the potential to largely replicate F2F interaction by creating three-dimensional, immersive digital environments where people can engage with one another in real-time (Hennig-Thurau et al., 2023). Unlike video conferencing, VR introduces avatar embodiment, spatial positioning, depth perception, and interactive elements, which may enhance social presence and create a more engaging and dynamic meeting experience.

VR simulates real-world environments using head-mounted displays, motion controllers, and spatial audio to create a sense of immersion. Avatars represent people in a VR meeting, allowing them to move, make gestures with their controllers, and interact within a shared space, thereby more closely mimicking F2F communication than 2D video platforms (Singh et al., 2022; Kimmel et al., 2024). Through real-time spatial audio, VR enables people to perceive sound directionally, allowing for more natural group discussions (Merz et al., 2024). Additionally, VR meeting platforms include interactive tools such as shared whiteboards and 3D object manipulation, resulting in a more engaging and collaborative experience (Qiu et al., 2023; Speidel et al., 2023).

By leveraging these features, VR has the potential to overcome the four limitations of MS Teams. First, VR environments enable users to communicate more effectively through gestures, posture, and facial expressions in real-time through avatars equipped with advanced motion tracking and facial recognition. These technologies allow nonverbal communication, such as nodding or smiling, which can enhance emotional connections in virtual interactions (Kimmel et al., 2024). Second, spatial awareness enables individuals to perceive their positions relative to others in a shared virtual environment, which can enhance engagement and conversational flow (Dean et al., 2014). Third, users in VR are isolated from external distractions through increased immersive involvement, increasing engagement (Baloian et al., 2023). Fourth, depth perception enhances the realism of virtual spaces, making it easier to collaborate on design, spatial tasks, or complex visual presentations (Rzayev et al., 2020).

Despite these advancements in VR, there is limited empirical evidence on how effectively VR replicates social presence compared to MS Teams and F2F meetings in real-world business settings. While VR theoretically offers a higher level of social presence than video conferencing, its effectiveness in real-world business meetings is uncertain. As the idea of a future metaverse gains increased attention, understanding how immersive technologies like VR influence engagement, collaboration, and communication becomes increasingly important. Organizations are starting to investigate how these environments might enhance remote teamwork, training, and engagement beyond what current video conferencing tools can provide. Therefore, to address this gap, this study aims to research the effectiveness of VR for mediated communication by comparing the perceived social presence of VR, F2F, and video conferencing in a business environment. This will contribute to the body of knowledge and gain practical insights for organizations interested in adopting immersive technologies to enhance virtual collaboration, engagement, and productivity.

This study addresses the following research question:

What is the difference in perceived social presence between virtual reality meetings, video conferencing, and face-to-face meetings in a business environment?

To address this research question, it is first necessary to establish a clear understanding of social presence and the factors that influence it. To achieve this, the following sub-questions are formulated:

1. What is social presence, and why is it important for effective communication in both physical and virtual meetings?
2. How does social presence vary across face-to-face, video conferencing, and VR?
3. Which factors influence social presence?
4. How can social presence be measured objectively to compare the meeting modalities?

1.1 Outline

This thesis is structured as follows: first, the systematic literature review details the literature search strategy, selection process, and thematic results, laying the foundation for the theory. Subsequently, the theory outlines the key theoretical concept of social presence, the factors influencing it across different meeting modalities, and how it is measured. This is followed by the research design, which describes the research objective and methodology. Next, the study's results are presented, followed by an analysis chapter that compares the findings to the literature. Lastly, the discussion and conclusion present the findings, outline the limitations, and offer suggestions for future research.

2 Systematic literature review

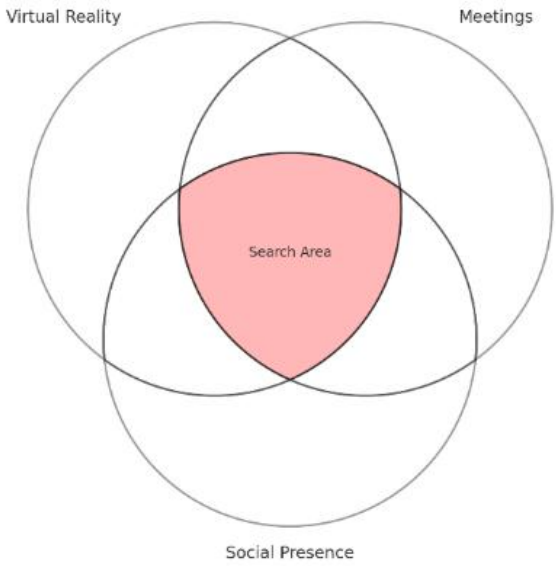
The PRISMA approach was utilized to conduct a literature review. PRISMA is the recognized standard for presenting evidence in systematic reviews and meta-analyses. The systematic literature review began with the primary research question.

2.1 Search strategy

Scopus and Web of Science are the selected databases that cover the majority of relevant papers available. The search constructs are 'social presence', 'virtual reality', and 'meetings'. 'Video conferencing' and 'face-to-face meetings' were intentionally excluded from the search string because the primary focus is on VR. Papers that solely discuss video conferencing or F2F meetings without the use of VR are not directly relevant to this research. However, comparative studies that analyze VR alongside these other modalities will naturally be included, as they incorporate VR as a key component. This approach ensures that the most relevant papers are captured to understand social presence in VR meetings while allowing meaningful comparisons with other modalities. The search string includes the constructs and their related terms and synonyms, as listed in the third column of Table 1. The search area encompasses the overlap between 'virtual reality', 'meetings', and 'social presence', as depicted in Figure 1. The search is conducted within each construct's article title, abstract, and keywords. Each search is documented, including the search date, the database used, the search terms or strings, and the number of results retrieved.

Figure 1

Literature search area. Overlap of VR, meetings, and social presence.



2.2 Selection process

The article selection process consists of inclusion and exclusion criteria, article screening, and extracting relevant data.

The inclusion criteria include peer-reviewed articles focusing on social presence, VR, and meetings. These also include articles that use synonyms or closely related terms, as listed in Table 1. The synonyms and closely related terms were identified by searching Google Scholar and utilizing Thesaurus. For example, it was found that researchers use the sense of presence, co-presence, or perceived presence interchangeably with social presence.

The exclusion criteria are articles not written in English, non-peer-reviewed articles, and articles on game design or prototype design. The articles retrieved from the search are screened based on title and abstract, and duplicates are removed. After that, the remaining articles are screened by full text and are assessed for eligibility using the inclusion and exclusion criteria. The number of articles after each stage is documented and can be found in the flow diagram depicted in Figure 2. Relevant data from the included articles are extracted into a standardized form.

Table 1

The used constructs with their related terms/synonyms and incorporated search string.

Construct	Related terms/synonyms	Search string
Social presence	Sense of presence, co-presence, perceived presence	"social presence" OR "sense of presence" OR "co-presence" OR "perceived presence"
Virtual reality	VR, immersive environments, XR, extended reality, 3D environments	"virtual reality" OR "VR" OR "immersive environments" OR "XR" OR "extended reality" OR "3D environments"
Meetings	Virtual meetings, online meetings, virtual collaboration, (tele)conferences, business meeting	"meetings" OR "virtual meetings" OR "online meetings" OR "virtual collaboration" OR "teleconferences" OR "conferences" OR "business meetings"

2.3 Results systematic literature review

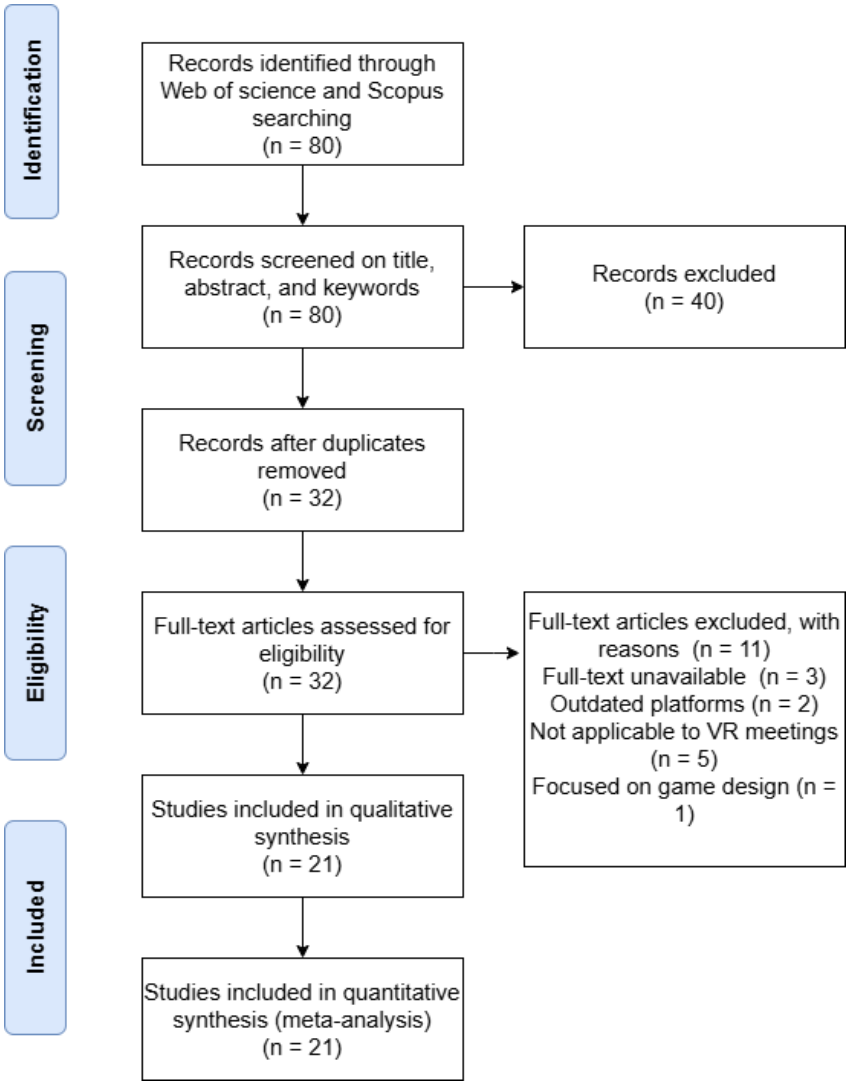
The results are now presented through a PRISMA flow diagram, which includes the selected papers, the trend of social presence in VR meetings, and the common themes identified within these papers.

2.3.1 PRISMA flow diagram

A total of 80 articles were retrieved from both databases. Of these 80, 40 were selected after reviewing the title, abstract, and keywords. Among the 40, there were eight duplicates, resulting in 32 articles being selected. Unfortunately, three articles had unavailable full texts and were thus removed from the set. The final collection consists of 21 articles after screening the full texts. Figure 2 illustrates the PRISMA flow diagram, depicting the flow of studies through each phase of the review process.

Figure 2

PRISMA flow diagram of the inclusion & exclusion process.



2.3.2 Papers included

The theory chapter includes 21 articles in Table 2, which presents the publication year and the number of citations. A few papers have zero citations; however, as these are mostly recent papers, it is appropriate to include them in the review.

Table 2

Papers included in the study.

Nr.	Author	Year	Citations
1	Adkins et al. (2024)	2024	0
2	Baloian et al. (2023)	2023	0
3	Chessa&Solari (2021)	2021	8
4	Cristina Gasch et al. (2024)	2024	0
5	Dean et al. (2014)	2014	7
6	Hennig-Thurau et al. (2023)	2023	174
7	Higgins et al. (2021)	2021	17
8	Lawrence et al. (2021)	2021	0
9	Merz et al. (2024)	2024	1
10	Michael Bonfert et al. (2023)	2023	11
11	Olt et al. (2024)	2024	1
12	Qiu et al. (2023)	2023	2
13	Rzayev et al. (2024)	2020	3
14	Simon Kimmel et al. (2024)	2024	0
15	Singh et al. (2022)	2022	4
16	Speidel et al. (2023)	2023	12
17	Steinicke et al. (2020)	2020	32
18	Sun Joo et al. (2021)	2021	42
19	Tüzün et al. (2019)	2019	8
20	Van Gent et al. (2024)	2024	0
21	Wang et al. (2024)	2024	2

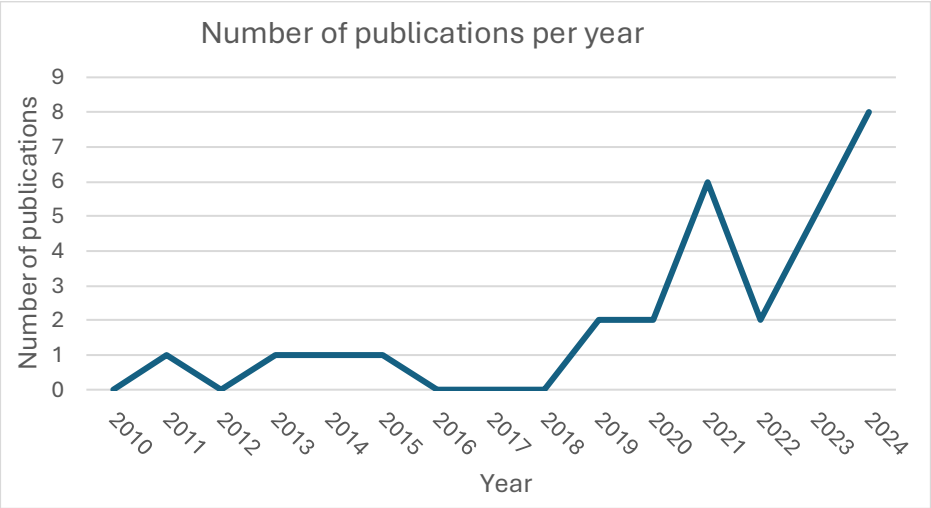
2.3.3 The trend of social presence in VR meetings

Each paper included in the study was categorized, and the number of publications per year was counted to create the graph shown in Figure 3. This figure illustrates the trends in publications from 2010 to 2024, encompassing the 21 articles included in the study. An upward trend, beginning in 2020, is observed in the annual number of publications on this research topic. Between 2010 and 2020, few publications were produced on this topic. This aligns with the

growing trend of hybrid working over the past five years. This could indicate that an increase in hybrid working has led to increased social presence research.

Figure 3

Number of publications per year.



Note. N = 21; an upward trend, starting in 2020, is observed.

2.3.4 Common themes

Table 3 presents the data categorization. Five common themes were identified across all papers in the systematic literature search:

- Conceptualizing Social Presence
- Factors Influencing Social Presence
- Social Presence in VR meetings
- Comparison across modalities
- Measuring Social Presence

The table presents the papers that addressed specific themes. These themes are the distinct sections discussed in the following chapter.

Table 3

Data categorization.

Paper	Conceptualizing Social Presence	Factors influencing Social Presence	Social Presence in VR meetings	Comparison across modalities	Measuring Social Presence
Adkins et al. (2024)					X
Baloian et al. (2023)		X			X
Chessa&Solari (2021)		X			
Cristina Gasch et al. (2024)		X			
Dean et al. (2014)		X			
Hennig-Thurau et al. (2023)	X				
Higgins et al. (2021)	X				
Lawrence et al. (2021)		X			
Merz et al. (2024)		X			
Michael Bonfert et al. (2023)			X		
Olt et al. (2024)				X	
Qiu et al. (2023)		X			
Rzayev et al. (2024)					X
Simon Kimmel et al. (2024)		X			
Singh et al. (2022)		X			
Speidel et al. (2023)			X		
Steinicke et al. (2020)				X	
Sun Joo et al. (2021)			X		
Tüzün et al. (2019)				X	
Van Gent et al. (2024)				X	
Wang et al. (2024)					X

3 Theory

This chapter is divided into five sections, each centered on a common theme outlined in Table 3. First, it conceptualizes social presence, followed by a discussion of the factors that influence it. Next, social presence in VR meetings is examined, followed by a comparison of social presence in F2F interactions, video conferencing, and VR. Finally, methods for measuring social presence are discussed, including the method utilized in this research.

3.1 Conceptualizing Social Presence

Social presence is a fundamental concept in understanding human communication and interaction across different modalities, whether F2F or mediated through technology. Short, Williams, and Christie (1976) first introduced the concept of social presence in telecommunication research, defining it as the degree to which a communication medium enables individuals to perceive others as being physically and socially present. In its earliest conceptualization, F2F interaction was considered the highest form of social presence, allowing for direct engagement, nonverbal communication, and shared spatial awareness. However, the increasing reliance on mediated communication has expanded this definition to include digital platforms, video conferencing, and immersive virtual environments (Biocca et al., 2003).

Social presence refers to the perceived sense of being with others in any given environment, whether physical or digital. The strength of social presence in a communication setting depends on how well a medium conveys interpersonal cues, emotional connection, and interaction fidelity (Biocca et al., 2003). For example, MS Teams and other video conferencing platforms rely on live video and real-time voice transmission to facilitate communication between remote users. In contrast, VR introduces spatial awareness, embodiment, and immersive engagement to create a stronger sense of shared space (Hennig-Thurau et al., 2023).

Thus, Hennig-Thurau et al. (2023) emphasize that social presence includes not just verbal communication but also spatial awareness, body language, and shared interactivity. Their research explores remote, mixed social interactions; digital experiences that closely mimic real-world social dynamics. They found that immersive VR creates a stronger sense of presence than video conferencing due to spatial and embodiment factors, although it still falls short of F2F interactions. Higgins et al. (2021) further contribute to this by focusing on the role of avatars and personalization in social presence. Their findings suggest that users feel a stronger connection and social engagement when they have personalized avatars that reflect their identity and expressions. This aligns with Biocca et al. (2003) theory of social presence, which suggests that the sense of being with others in digital environments is influenced not

only by the communication medium but also by the extent to which individuals feel represented in the space.

In summary, social presence is a key aspect of human communication that applies to both physical and digital environments. While F2F remains the benchmark for achieving maximum social presence, mediated interactions, such as video conferencing and VR, provide alternative ways to establish presence remotely. The extent to which these platforms successfully replicate the perception of being with others depends on their ability to facilitate interpersonal cues, spatial awareness, and embodied engagement.

3.2 Factors Influencing Social Presence

Social presence is determined by various factors that either enhance or inhibit the feeling of being "with others" in shared environments, whether F2F, video conferencing, or VR. These factors can be categorized into user representation and embodiment, cognitive and emotional engagement, device and tool integration, and the impact of the meeting environment. These are described in the following sections.

3.2.1 User representation and embodiment

User representation is a key factor in social presence within digital environments. In F2F meetings, user representation is naturally strong because individuals interact with their full bodies, facial expressions, and gestures in real-time. In contrast, video conferencing platforms like MS Teams limit embodiment to a flat screen, where only faces and upper bodies are visible. While video feeds allow for some nonverbal cues, they fail to fully capture body language, spatial movement, and physical engagement. VR attempts to bridge this gap by using avatars to represent users. The level of realism and motion-tracking accuracy of these avatars is crucial for achieving a high level of social presence. Dean et al. (2014) showed that mapping real-world movements to avatars significantly increases immersion and engagement. Their findings show that facial expression tracking and realistic gestures enhance the perceived social presence, as users can more naturally see each other's emotions. Similarly, Simon Kimmel et al. (2024) found that accurately replicating body movement increases social presence in collaborative VR environments. When kinetic accuracy improves, users perceive that their interactions more closely resemble real-life communication. Additionally, Higgins et al. (2021) looked into the role of avatar personalization in enhancing social presence. Their study showed that users experience greater emotional connection and engagement when avatars reflect their personal identity.

In summary, people feel more naturally represented in F2F interactions, whereas MS Teams lacks this sense of embodiment. In contrast, VR enables embodied interaction through avatars, which can enhance or diminish social presence depending on the quality and responsiveness of the avatars.

3.2.2 Cognitive and emotional engagement

Next to user representation and embodiment, social presence is also influenced by the extent to which a communication medium demands cognitive and emotional engagement from its users. In F2F meetings, engagement is naturally high because direct physical interaction demands attention, social norms discourage multitasking, and environmental cues keep participants mentally focused. However, video conferencing often leads to passive participation, as users can turn off their cameras, multitask, or disengage without being noticed. VR requires high cognitive engagement because it immerses users in a three-dimensional, interactive space. Baloian et al. (2023) employed EEG analysis to demonstrate that VR meetings necessitate greater cognitive engagement than video conferencing. The results indicate that higher engagement is correlated with stronger perceptions of social presence, as immersion encourages users to interact more attentively. However, this increased engagement can also lead to cognitive overload, particularly for users unfamiliar with VR environments.

Another key aspect of engagement is emotional fatigue. Speidel et al. (2023) investigated the impact of fatigue on social presence across various modalities. Their study found that VR mitigates the effects of "Zoom fatigue" often experienced in MS Teams meetings, leading to higher levels of participation and presence in virtual spaces. While F2F meetings rarely cause significant fatigue due to natural human interaction dynamics, both MS Teams and VR require cognitive effort to interpret digital representations of presence, which affects long-term engagement and interaction quality.

3.2.3 Device inclusivity and interaction tools

The ability of users to access and interact within a given medium also influences social presence. In F2F meetings, technology is not a limiting factor; physical presence is all that is required. However, in video conferencing, social presence is influenced by factors such as video quality, microphone clarity, and internet stability. Poor audiovisual conditions reduce the ability to perceive others clearly, limiting engagement. In VR, hardware and input devices have a significant impact on social presence. Merz et al. (2024) investigated the impact of different VR devices (e.g., standalone headsets vs. PC-based VR systems) on user perception. Their findings indicate that while hardware variations influenced self-perception, they did not

have a significant impact on social presence or collaboration quality. This suggests that VR inclusivity is possible, meaning different hardware setups can still have meaningful social interactions. Additionally, Qiu et al. (2023) investigated hybrid virtual meetings, where some participants utilized VR while others participated via traditional video conferencing. Their research found that hybrid setups work best when avatar and environmental consistency is maintained across platforms. This demonstrates that even in mixed environments, social presence can be optimized by ensuring seamless integration across different modalities. Finally, Lawrence et al. (2021) studied how high-fidelity telepresence improves social presence compared to standard 2D video conferencing. Their study showed that higher audiovisual fidelity enhances copresence, engagement, and nonverbal communication. This aligns with broader findings that spatial audio, natural voice rendering, and realistic environmental details help replicate the experience of F2F meetings in VR.

3.2.4 The impact of the meeting environment on social presence

The environment in which a meeting takes place, whether it's physical or digital, can significantly influence how social presence is experienced. In F2F settings, presence is natural because people are physically present in shared spaces, allowing for spontaneous interactions, spatial awareness, and environmental context. The physical environment sets expectations for interaction, with seating arrangements, room acoustics, and body positioning contributing to the sense of co-presence. However, in video conferencing and VR, the digital environment must be intentionally designed to compensate for the lack of physicality, helping to maintain engagement and social connection.

Merz et al. (2024) indirectly support this by showing that device characteristics impact self-perception and collaboration, suggesting that the surrounding digital space may also influence presence. How people experience a virtual meeting room, whether it closely represents their traditional office space, for example, can affect their sense of co-presence. Qiu et al. (2023) found that consistent platform environments across different devices helped maintain social presence, implying that a well-structured digital environment can increase immersion and engagement across different meeting modalities. Additionally, high-fidelity telepresence systems, such as those examined by Lawrence et al. (2021), have been shown to enhance presence through realistic audiovisual settings. Their study found that spatial audio, realistic visual depth, and interactive environmental elements contribute to a stronger sense of shared space, making digital meetings feel more like physical meetings. This suggests that features such as spatial sound, realistic backgrounds, and interactive tools like whiteboards or

shared screens can have a positive impact on the sense of presence. However, if the VR environment is too abstract, cluttered, or unrealistic, it may lower the sense of social presence rather than enhance it.

In summary, social presence is influenced by various factors that differ across F2F, MS Teams, and VR meetings. While F2F interactions provide the most perceived social presence, digital platforms attempt to replicate these experiences. By optimizing avatar realism, reducing fatigue, improving hardware accessibility, and designing immersive meeting environments, VR may become a more viable alternative for remote collaboration.

3.3 Social presence in VR meetings

VR enables users to interact in a shared virtual space using avatars, spatial audio, and dynamic environments. These elements are designed to replicate the natural dynamics of F2F interactions. They together create stronger social connections, engagement, and interactivity. The extent to which VR enhances social presence depends on spatial awareness, avatar realism, and interactive capabilities.

3.3.1 Spatial awareness

VR is seen as an effective medium for social presence due to its ability to generate a stronger sense of co-presence, making users feel as though they are physically together in a shared space, having spatial awareness of each other. Michael Bonfert et al. (2023) examined the effectiveness of social presence in VR meetings. They found that VR platforms facilitate a greater sense of immersion and interaction through spatial positioning and real-time movement tracking. Participants reported feeling more connected to their virtual counterparts due to the ability to move within a 3D space, experience directional audio, and engage in natural conversations. Next, the structure of VR meetings themselves can influence social presence. Sun Joo et al. (2021) studied social presence in virtual academic conferences, finding that VR enabled dynamic and spontaneous interactions, which are often missing in other digital communication formats. Participants in the study noted that spatial proximity and the ability to navigate virtual rooms improved their sense of being part of a larger event, increasing their engagement and co-presence. However, the study also highlighted challenges, including usability barriers, learning curves, and discomfort with VR hardware, which can sometimes detract from social presence. Speidel et al. (2023) examined the impact of VR on meeting fatigue, finding that while VR can reduce some of the cognitive drain associated with traditional video meetings, it also introduces new forms of fatigue due to the demands of spatial

awareness, prolonged headset use, and cognitive overload. Therefore, VR meetings must be structured thoughtfully to balance immersion with user comfort.

3.3.2 Avatar realism

The realism of avatar representation is important for determining how users experience social presence in VR environments. Singh et al. (2022) conducted a study on avatar realism in remote business meetings, showing that high-fidelity avatars correlated with increased social presence and engagement. Their findings suggest that users feel more present in the virtual space when their avatars are expressive and interactive, reinforcing the idea that realistic body movements and facial tracking contribute to deeper social interactions. Another important element of avatar realism is gaze and proxemics, which refer to how users perceive eye contact, body positioning, and spatial relationships in virtual environments. Wang et al. (2024) investigated the role of gaze alignment and proximity adjustments in VR meetings, demonstrating that improving eye contact and spatial positioning enhances mutual attention and co-presence. Their research revealed that the design of a VR space should facilitate realistic eye movements and spatial interactions to improve engagement among participants.

3.3.3 Interactive capabilities

The ability to interact with the virtual environment and other participants meaningfully is another aspect of increasing social presence in VR. VR enables users to manipulate objects, use interactive whiteboards, and engage in real-time collaboration within the virtual space. Qiu et al. (2023) studied hybrid virtual meetings, finding that the presence of interactive tools, such as shared 3D objects and gesture-based interactions, significantly improves engagement and co-presence. Their research suggests that VR meetings feel more natural when users can physically interact with their environment instead of passively observing avatars in a digital space. Additionally, Merz et al. (2024) found that intuitive controls and user interaction enhance social presence through device integration and interaction tools. These findings demonstrate that interactive features are crucial for enabling participants to naturally engage with one another and their surroundings, much like they would in a real-life setting.

3.4 Comparison across modalities

The differences in social presence across F2F, video conferencing, and VR have been partially covered. Social presence has been found to be the highest for F2F, followed by VR and video conferencing. It is important to understand how these modalities compare to determine whether VR meetings can effectively replicate F2F communication and surpass video conferencing.

3.4.1 Differences in Social Presence across modalities

Several studies have directly compared social presence across these modalities. Van Gent et al. (2024) conducted an experimental comparison of F2F, video conferencing, and VR meetings, finding that while VR meetings were perceived as more enjoyable, task engagement remained similar across all three modalities. This suggests that while VR enhances social presence, it does not automatically translate to higher engagement or productivity. Similarly, Steinicke et al. (2020) found that VR meetings outperformed video conferencing in terms of social presence and immersion, indicating that VR is a more engaging alternative to traditional digital meetings. Further supporting this comparison, Hennig-Thurau et al. (2023) provided a theoretical and empirical framework for social presence in digital communication. Their research showed that social presence is strongest in F2F settings, followed by VR, with video conferencing ranking lowest. They argue that VR's advantage lies in its ability to replicate spatial and nonverbal elements missing from video conferencing. However, they also note that VR has usability challenges that F2F and video conferencing do not, such as headset discomfort, motion sickness, and cognitive load.

3.4.2 Emotional connection and fatigue across modalities

Another important factor in comparing social presence across modalities is how emotionally connected and fatigued participants feel during interactions. Speidel et al. (2023) investigated the impact of VR on Zoom fatigue, finding that VR meetings can help mitigate the fatigue commonly associated with video conferencing by making interactions feel more immersive and dynamic. However, they also noted that VR meetings introduce new types of fatigue, including cognitive overload and discomfort from prolonged headset use. Additionally, Chessa and Solari (2021) found that VR enhances social presence compared to video conferencing; however, usability challenges remain a barrier to widespread adoption. This highlights the idea that while VR improves certain aspects of social presence, it does not completely eliminate the limitations of mediated communication. This is because, in F2F meetings, emotional connection is naturally sustained through direct engagement, body language, and real-time feedback. In contrast, both video conferencing and VR require technological enhancements to replicate these effects.

Comparing social presence across F2F, video conferencing, and VR highlights the trade-offs of each modality. F2F remains the strongest medium for social presence, as it provides full physical interaction, spontaneous communication, and natural emotional connection. Video conferencing is effective, but it suffers from limitations in embodiment and

spatial presence, which can make social interactions feel more constrained. VR, on the other hand, bridges the gap between video conferencing and F2F by offering avatar embodiment, spatial awareness, and immersive engagement. However, it introduces new usability challenges, such as cognitive load and hardware limitations, that need to be addressed for VR to become a fully viable alternative to traditional meetings.

3.5 Measuring Social Presence

The concept of social presence and its relationship to face-to-face (F2F), video conferencing, and virtual reality (VR) has been outlined. The following section explains how social presence can be measured and the method used in the experiment. Social presence has been measured using a variety of approaches, which can be categorized into subjective self-report measures, behavioral observations, and physiological measures.

3.5.1 Self-report measures: The Networked Minds Social Presence Inventory (NMSPI)

Self-report questionnaires are the most used method for evaluating social presence; they allow users to directly express their perceptions of co-presence, engagement, and interaction quality. One of the most used instruments is the Networked Minds Social Presence Inventory (NMSPI), developed by Harms and Biocca (2004). The NMSPI is designed to measure social presence across six dimensions:

1. Co-presence: the degree to which participants perceive themselves as physically present in the same shared space or environment.
2. Attentional allocation: the amount of attention participants devote to others during interactions.
3. Perceived message understanding: participants' perception of how clearly messages are exchanged and understood during communication.
4. Perceived affective understanding: the accuracy with which participants perceive and interpret others' emotional states during interactions.
5. Perceived emotional interdependence: the extent to which participants feel their emotions affect, or are affected by, the emotions of others, emphasizing emotional connectivity.
6. Perceived behavioral interdependence: the degree to which participants perceive their behaviors as influencing, or being influenced by, the behaviors of their team members.

The NMSPI is particularly useful for comparing social presence across different modalities. This is because the NMSPI focuses on how people perceive their social interactions, the questions are applicable to each modality.

3.5.2 Behavioral and physiological approaches to measuring social presence

Behavioral and physiological approaches offer alternative methods for assessing social presence. Behavioral tracking, for example, analyzes eye contact, spatial positioning, and interaction patterns to determine how users engage in a digital environment. Wang et al. (2024) found that gaze alignment and proxemic transformations in VR meetings significantly impact social presence, showing that users' spatial positioning and visual attention can be indicators of social presence. Similarly, Qiu et al. (2023) examined interaction consistency across hybrid meetings (VR vs. video conferencing), highlighting how social behaviors differ across modalities.

Physiological measures, such as electroencephalography (EEG), eye tracking, and heart rate variability (HRV), provide objective insights into cognitive engagement and emotional responses to mediated presence. Baloian et al. (2023) employed EEG analysis to compare cognitive load in VR and video conferencing, revealing that VR meetings necessitated higher cognitive engagement but also elicited greater perceptions of social presence. However, while physiological measures provide valuable neuroscientific evidence of social presence, they require specialized equipment and controlled lab environments, making them less feasible for this study.

Since this study focuses on comparing social presence across different modalities, the NMSPI questionnaire is used, along with qualitative behavioral observations recorded through an observation checklist. This approach ensures that the study captures both perceived and interactional measures of social presence.

3.6 From theory to experiment

This section summarizes how the theoretical insights from the literature review are applied in the experimental design of this study. Table 4 outlines key concepts of social presence, engagement, and collaboration in virtual meetings and explains how they are integrated into the research design. This ensures that the study is grounded in the Literature while allowing for systematic evaluation of social presence across different meeting modalities.

Table 4

Key Insights from the Literature and Their Application in the Research Design.

Theoretical concept	Key insights from the literature	Application in research design
Social Presence	Social presence refers to the sense of "being with others" in a mediated communication setting (Biocca et al., 2003). VR enhances social presence through spatial awareness, avatar embodiment, and immersive engagement (Hennig-Thurau et al., 2023).	The study compares perceived social presence across three meeting modalities: VR, MS Teams, and F2F. The NMSPI questionnaire is used to measure participants' subjective experiences of social presence.
Comparison across modalities	F2F interactions provide the highest social presence, followed by VR, then video conferencing (Van Gent et al., 2024; Steinicke et al., 2020). VR enhances immersion and engagement, but introduces usability challenges, such as motion sickness and cognitive load (Speidel et al., 2023).	The experiment tests all three modalities under controlled conditions to analyze how social presence differs in VR, MS Teams, and F2F. Observations capture the quality of engagement, collaboration, and interaction to supplement questionnaire data.
Factors Influencing Social Presence	Social presence is realized by avatar realism, cognitive engagement, spatial awareness, and usability (Dean et al., 2014; Singh et al., 2022; Wang et al., 2024). Avatars in VR with motion tracking and facial expressions enhance engagement, but they are still not perfect representations of reality.	Participants experience VR using Spatial.io with avatars and spatial audio, allowing for comparison to MS Teams (which lacks embodiment) and F2F (which has a full physical presence). Observations focus on body language, interaction frequency, and engagement levels.

Measuring social presence	Social presence is measured through self-report scales, behavioral observations, and physiological measures (Biocca et al., 2003). The Networked Minds Social Presence Inventory (NMSPI) is a validated tool to assess perceived social presence.	The study uses the NMSPI questionnaire after each meeting modality to measure co-presence, psychological involvement, and behavioral engagement. Observations provide qualitative validation of social presence indicators.
Engagement and collaboration in VR	VR meetings enhance collaboration through spatial awareness and shared 3D workspaces (Bonfert et al., 2023; Qiu et al., 2023). However, technical difficulties and learning curves can reduce effectiveness (Speidel et al., 2023).	Engagement and collaboration are assessed through observations of verbal participation, eye contact, and tool usage (e.g., whiteboards in VR). Data from NMSPI and observational analysis will help evaluate if VR meetings facilitate more effective collaboration than MS Teams.

4 Methodology

This chapter outlines the research objective and describes the research design, including the study's phases, data collection methods, and analysis procedures. It also addresses the reliability and validity of the study.

4.1 Research objective

The primary objective of this study is to investigate how social presence in VR meetings compares to MS Teams and F2F meetings in a business environment. Current research has identified limitations in video conferencing tools related to social presence; however, empirical evidence on how effectively VR can overcome these limitations remains limited. By comparing social presence across these three modalities, this research aims to fill this gap by providing empirical data on the potential of VR. The insights gained from this study can offer practical implications for organizations considering the adoption of immersive technologies for hybrid working scenarios.

4.2 Research Design

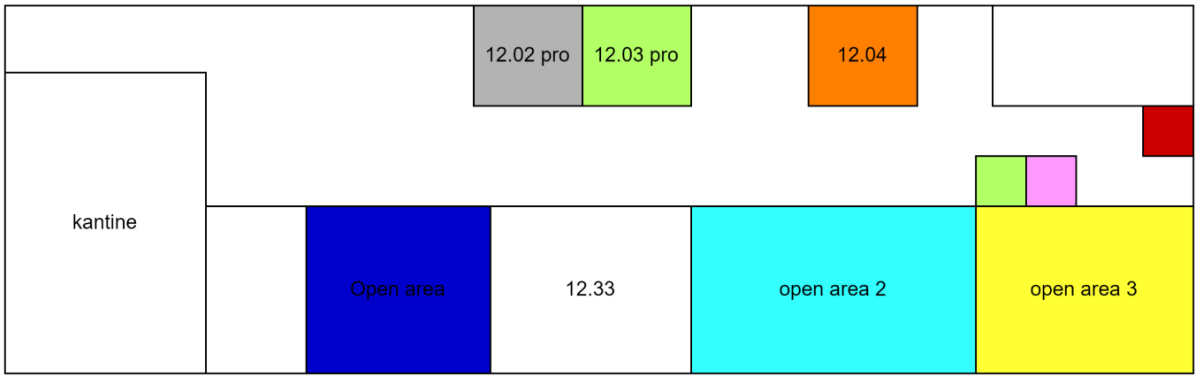
This research employed a mixed-methods approach. An experimental design was combined with qualitative observations. The following sections outline the study context, participants, and equipment.

4.2.1 Study context

The study took place within a business environment at CGI, specifically within the Retail Team in the Netherlands. All three research activities (training, pilot VR meeting, and experiment) took place at the CGI office in Amsterdam. The office contained multiple meeting rooms and two phone booths, which were used to isolate each participant during the VR and MS Teams meetings. A floor map was created of the office, where each location was assigned a color. During the training, each participant was assigned a color that corresponded to the color of the VR headset they used. Each VR headset was placed on the designated location on the map before the start of the meeting. The map is depicted in Figure 4.

Figure 4

Floor map of the CGI office with color distribution. Each color corresponds to the marking on the VR headset that was located in that area.



4.2.2 Participants

Participants were business professionals who regularly used MS Teams for their meetings. A total of seven participants participated in the study. The participants represented a diverse range of roles within CGI, from junior consultants to director consultants. While all participants were experienced with MS Teams, their familiarity with VR technology varied.

4.2.3 Equipment

The University of Twente supplied the VR headsets required for the study. A total of seven VR headsets were used. These consisted of 1 Meta Quest Pro device and 6 Meta Quest 2 devices.

4.3 Study setup

The study consisted of three phases, which are described in the following sections.

1. Training & avatar creation
2. Pilot VR meeting
3. Experiment

4.3.1 Training & avatar creation

Every participant received VR training and created their own avatar prior to the experiment and the pilot VR meeting. This ensured that they could effectively participate during the meeting experience.

Avatar creation

The Spatial platform was utilized for the pilot VR meeting and experiment, which will be explained in more detail later. It allowed users to create avatars by taking photos of themselves

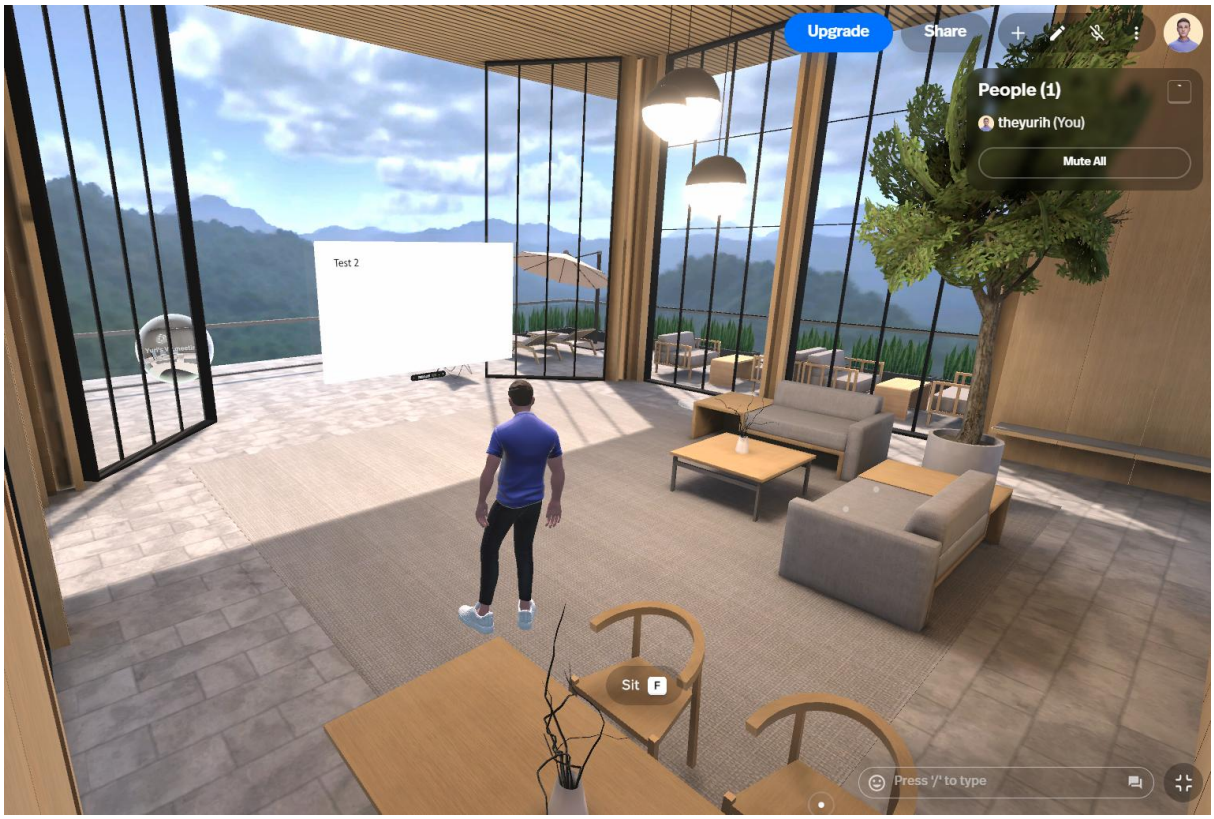
and using these photographs to generate their corresponding avatars. A guide was created and sent to each participant, explaining how to set up an account and create an avatar on Spatial.

Training

Participants were trained one-on-one before the pilot VR meeting. The participants were instructed on how to wear the VR headset, navigate the virtual space, and use basic controls, including movement, object selection, interaction with other avatars, and how to sit down or stand up. This allowed them to become comfortable with the equipment, ask questions, and troubleshoot any issues in advance. Every participant received the same training. The reason for the training was that participants might have struggled with navigating the VR environment, which could have led to frustration and distraction during the VR meeting. This could have potentially skewed the results because participants might have attributed negative experiences to the VR platform when they were actually unfamiliar with the technology. The training ensured that participants who encountered challenges during the VR meeting did so due to issues with platform usability. The training was conducted in a Spatial environment, as depicted in Figure 5. The training was conducted by screen-sharing the VR headset with the researcher's laptop. This allowed him to see what the participant saw and give input accordingly.

Figure 5

Training environment in Spatial. This environment was used during the training, where each participant was taught to move, sit down/stand up, and work with the PowerPoint presentation.



4.3.2 Pilot VR meeting

The pilot VR meeting was used as a test before the actual experiment. This meeting was conducted entirely in VR to test the VR features and identify any issues. By doing a try-out VR meeting, each participant also experienced VR for the first time. Experiencing a new technology for the first time is always fun and interesting, and this effect could have influenced the experiment's results. Therefore, a try-out was done first to ensure that the experiment would be the second time they experienced VR. The setup of the meeting and the meeting platform are explained in the following sections.

Meeting setup

The retail team at CGI holds a team meeting every six weeks. This meeting typically takes place in the evening after work hours, as each member is usually with a client during the day. It can be conducted either via MS Teams or face-to-face. The agenda typically begins with updates on business performance, followed by a more interactive segment, such as small-group brainstorming sessions to improve KPIs. The team leader ensures that the meeting remains

relevant for every team member, actively soliciting and receiving feedback. This focus on relevance also boosts attendance among team members.

The pilot VR meeting was conducted during one of the team meeting slots. The pilot VR meeting took place at the CGI office in Amsterdam. A total of seven team members were present, all of whom wore VR headsets. To evaluate the platform's spatial audio functionality, each participant was seated in a separate room to ensure that they could not hear one another without the use of a microphone.

The meeting agenda resembled that of a typical team meeting. The team leader began by sharing insights on business performance, followed by a brainstorming session. Initially, the entire group was seated in a virtual meeting room equipped with a PowerPoint screen and a large table with eight chairs. During this time, the team leader delivered his PowerPoint presentation. After the introductory segment, the group was divided into smaller teams of 2 to 3 members. Each group then moved to a separate area within the meeting room. They had 5-10 minutes to brainstorm ways to improve KPIs. After the brainstorming session, each group reconvened in the virtual meeting room. Then they were invited to present their ideas to the rest of the team, fostering an interactive experience for all participants. Members engaged with each other by asking questions and discussing the points raised. The meeting concluded with the team leader summarizing the key points discussed.

4.3.3 Meeting platform

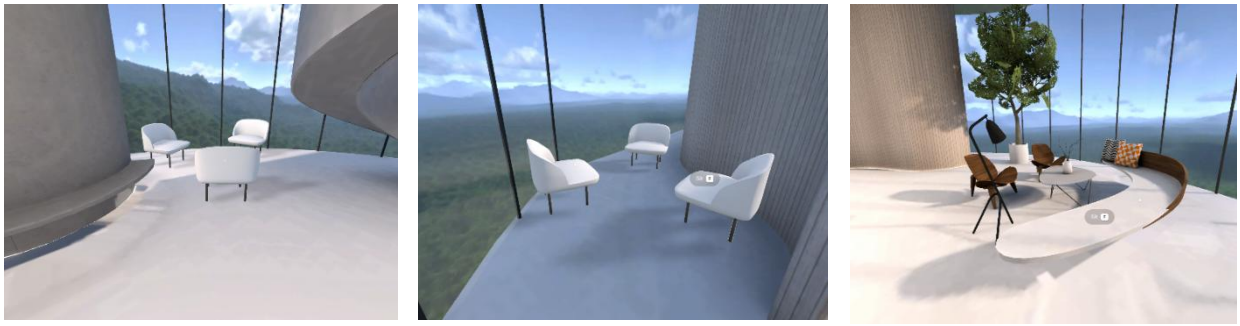
Spatial was used as the VR meeting platform. Spatial is a widely recognized virtual collaboration tool designed to create immersive meeting environments. It enables users to interact with customizable avatars that feature motion tracking, spatial audio, and interactive 3D objects. The platform supports real-time collaboration through features like virtual whiteboards, document sharing, and breakout rooms. For the meeting environment, one of the existing premium templates was used, which resembled a meeting room, as depicted in Figure 6. It featured a large table with chairs surrounding it, a whiteboard displaying a PowerPoint presentation, and multiple areas to direct people to. The room was adapted by adding extra chairs used for the breakout session. The break-out locations are depicted in Figure 7.

Figure 6

The VR meeting room environment used for the Pilot and experiment VR meeting.

Figure 7

Break-out locations used during the Pilot VR meeting.



4.3.4 Experiment

The experiment meeting involved the same group of seven participants, who experienced three different meeting modalities:

1. VR
2. MS Teams
3. F2F

The experiment, similar to the pilot VR meeting, also took place during one of the team meeting slots. The meeting lasted 60 minutes, starting with a 5-minute introduction and an explanation of the setup in a real-life meeting room. Next, the participants were moved to their designated

location, where VR headsets and laptops were prepared. The participants first met in VR for 15 minutes in the same environment used during the pilot VR meeting. One participant was asked to prepare a work-related topic and lead a discussion on it. After 15 minutes, each participant completed the NMSPI questionnaire regarding their VR experience. Following this, they met for another 15 minutes in MS Teams using their laptops, during which one participant created a discussion on a work-related topic. The NMSPI questionnaire was completed once more. Subsequently, all participants moved to the real-life meeting room to meet for 15 minutes and discuss another work-related topic. Afterward, they completed the NMSPI questionnaire based on their F2F experience.

4.4 Data collection

Data were collected through the NMSPI questionnaires and observations during the experiment, which are now explained in more detail.

4.4.1 NMSPI-questionnaire

The Networked Minds Social Presence Inventory (NMSPI) was used to measure social presence across three meeting modalities: VR, MS Teams, and F2F. After each meeting session, participants completed the NMSPI questionnaire, which consisted of a series of 5 Likert-scale questions designed to assess different dimensions of social presence. Participants completed the NMSPI questionnaire after each meeting to systematically compare differences in social presence between VR, MS Teams, and F2F interactions.

The NMSPI focused on six dimensions of social presence, which were explained in the "Measuring Social Presence" section of the theory. The original questionnaire consisted of 36 questions. It was decided to reduce the number of questions to 18, 3 per dimension, to decrease the time required for participants to complete the questionnaire. The choice of questions to retain was made based on the factor loadings provided in the paper. The questions with the highest factor loadings were retained. Two control questions were added, which were used in the quantitative analysis to verify the questionnaire results, which are:

1. The degree of perceived social presence (grade 1-10)
2. The effectiveness of the meeting platform for this meeting purpose (grade 1-10)

The questionnaire can be found in Appendix F.

4.4.2 Observation during the experiment

The researcher and an assistant were present during the experiment to ensure the smooth operation of the VR equipment and to assist participants with any technical issues. Both the assistant and researcher observed participant behavior to gather observational data on their experiences with the VR platform. The researcher and assistant observed how participants used their avatars; for example, they noted whether participants actively used gestures by moving their controllers during conversations. The session was also recorded to allow for double-checking of observations. To ensure consistency, the researcher used an observation checklist to track behaviors and events related to the focus areas systematically. The checklist, found in Appendix G, included predefined criteria, such as signs of engagement and collaboration. Other observational notes on noteworthy behaviors of the participants were also recorded.

4.5 Data analysis

The data collected from the NMSPI questionnaires and observations were analyzed using a combination of quantitative and qualitative methods. This approach ensured a comprehensive understanding of how social presence differed across VR, MS Teams, and F2F meetings.

4.5.1 NMSPI-Questionnaire Analysis

The NMSPI questionnaire responses were analyzed using descriptive and inferential statistical methods. First, descriptive statistics were used to calculate the mean and standard deviation for each of the 6 dimensions of social presence across the three meeting conditions. Then, a Repeated-Measures ANOVA was conducted to determine if there were statistically significant differences in social presence scores between VR, MS Teams, and F2F meetings. If significant differences were found, post hoc pairwise t-tests were used to identify which specific modalities differed significantly from one another. This statistical approach enabled a quantitative evaluation of whether VR meetings fostered a stronger sense of social presence compared to video conferencing and how both methods compared to F2F meetings. Linear mixed models were then conducted to determine whether the differences persisted after accounting for participants' individual differences. Finally, a Repeated-Measures ANOVA was conducted for each dimension of social presence to show which dimensions differed the most between the three modalities.

4.5.2 Observational Data Analysis

The observational data were analyzed using the structured notes collected through the observation checklist. Behaviors such as head turning, leaning forward, gesturing, and

indicators of engagement levels were documented for each modality. These observations were then summarized descriptively to identify noticeable differences or similarities in participant behaviors across the three meeting conditions. Attention was given to non-verbal interactions, engagement cues, collaboration, and technical issues. The findings from these observations were subsequently compared across meeting modalities to provide insights into the factors influencing perceived social presence.

4.6 Reliability and Validity

This section describes the steps taken to ensure the reliability and validity of this research.

4.6.1 Reliability

Reliability in research refers to consistency and replicability, meaning that the study should have yielded similar results if repeated under the same conditions. To ensure reliability, standardized procedures were implemented at all stages of the study.

First, structured data collection methods ensured consistency across all participants. The NMSPI questionnaire provided a systematic and predefined approach to measuring social presence, ensuring that all participants assessed their experiences using the same criteria. Similarly, observations were documented using a standardized checklist, which helped to systematically track behaviors such as engagement, interaction, and technical challenges across all meeting modalities.

Second, controlled experimental conditions further enhanced reliability. The study followed a fixed sequence of meeting modalities (VR, MS Teams, F2F) and identical meeting tasks, ensuring that any observed differences in social presence came from the medium itself rather than variations in task structure or meeting flow. Additionally, participants received VR training, ensuring that all individuals had a basic understanding of VR functionality, thereby reducing the likelihood that unfamiliarity with the technology influenced the results. Participants also took part in a VR test meeting, removing the first-time experience effect.

Lastly, the researcher followed a predefined observational framework that systematically recorded nonverbal behaviors, engagement levels, and collaboration patterns. This minimized subjective interpretation and ensured that all participants were assessed under the same criteria. These measures together ensured that observed differences in engagement, collaboration, and social presence reflected genuine differences between modalities rather than inconsistencies in the research process.

4.6.2 Validity

Validity refers to the accuracy and applicability of the study's findings. Several steps had been taken to ensure that the results were robust, unbiased, and accurately reflected participants' experiences.

To increase internal validity, this study employed triangulation by collecting data from multiple sources, including quantitative survey responses (NMSPI scores) and qualitative observational data. This ensured that both self-reported participant experiences and researcher-recorded behavioral indicators supported conclusions about social presence. By cross-referencing these data points, the study reduced the risk of misinterpretation and enhanced confidence in its findings.

To minimize researcher bias and enhance confirmability, data collection was conducted in accordance with strict procedures. The NMSPI questionnaire ensured structured and standardized responses, thereby reducing the influence of personal interpretation. Observations were recorded using a predefined checklist, ensuring that behavioral indicators were assessed objectively and consistently. Additionally, results were reviewed with the thesis supervisor, who provided external validation and helped reduce potential bias in the analysis and interpretation. Dependability was achieved through transparent documentation of all research steps, including data collection, analysis procedures, and coding techniques. This ensured that the study could be replicated by future researchers, strengthening the reliability of the findings.

Lastly, transferability was ensured by providing a detailed description of the study's context, including the participant demographics, meeting structure, and experimental setup. This description allowed researchers and professionals in similar fields to determine whether the study's findings could be applied to their own business, academic, or virtual collaboration settings.

5 Results

This chapter presents the study's findings, structured into qualitative and quantitative analyses, which compare participant experiences of social presence and interactions in VR, MS Teams, and F2F meetings. The qualitative section uses the observations across four themes: non-verbal communication, engagement, collaboration, and technical issues. The quantitative section uses Repeated Measures ANOVA and Linear Mixed Models to statistically evaluate the effects of meeting type on participant ratings, perceived social presence, and meeting effectiveness.

5.1 Qualitative analysis

This section presents the qualitative findings from the experiment's observations. The observations focused on four aspects of meeting dynamics: non-verbal cues, engagement, collaboration, and technical issues.

The observations will now be analyzed to compare participant behavior and interaction across the modalities. The complete list of observations is presented in Appendix E. This section begins with a detailed discussion of each theme and ends with a summarization table.

5.1.1 Method of analysis

The observational data were analyzed using the structured notes collected through the observation checklist. Behaviors such as head turning, leaning forward, gesturing, and indicators of engagement levels were documented for each modality. These observations were then summarized descriptively to identify noticeable differences or similarities in participant behaviors across the three meeting conditions. Attention was given to non-verbal interactions, engagement cues, collaboration, and technical issues. The findings from these observations were subsequently compared across meeting modalities to provide insights into the factors influencing perceived social presence.

5.1.2 Non-verbal communication

Observations revealed distinct differences in nonverbal communication across the meeting types. In VR, participants showed limited head movement and used minimal gesturing. The active speaker frequently moved their arms, but these gestures lacked clear communicative intent. Participants turned their heads toward the speaker or the PowerPoint presentation. In MS Teams, three participants were not consistently looking at the screen, and hand gestures were not used much. When hand gestures were used, they were not visible due to the limited camera frame. Some participants displayed engagement through nodding, while others slouched, signaling disengagement. One observer noted a lack of visible interest among

participants. In contrast, the F2F meeting demonstrated the most natural non-verbal interactions. Participants frequently switched between looking at the speaker and the PowerPoint presentation, and eye contact was more consistently maintained. Gestures were actively used, particularly by speakers, and nodding was a typical response when eye contact was made.

5.1.3 Engagement

Engagement varied across the three meeting types. In VR, participants initially struggled with setting up their positions, such as finding their chairs, which distracted them from the presentation. However, once settled, participants remained engaged, with no multitasking observed. The meeting was dominated by a single speaker, with limited instances of verbal interaction from the other group members. In MS Teams meetings, engagement was inconsistent. While an active speaker was present, observers noted instances of inattentiveness, as some participants looked away from their screens or engaged in multitasking. Some questions were asked, indicating occasional participation; however, overall engagement was lower than in F2F meetings. F2F meetings demonstrated the highest engagement levels. Most participants actively contributed, and the presence of two speakers facilitated a more interactive discussion. There were also moments of laughter and social interaction, which were not observed in VR or MS Teams, suggesting that F2F meetings are more dynamic and engaging.

5.1.4 Collaboration

Collaboration and turn-taking varied across meeting types. In VR and MS Teams, turn-taking was largely absent, with a single speaker dominating the session. Participants in these modalities primarily listened rather than engage in a dynamic conversation. While the active speaker in MS Teams sometimes addressed participants by name, this did not lead to increased interaction or structured dialogue. In contrast, F2F meetings showed a smoother flow. Participants took turns speaking, asked each other for additional information, and actively assisted one another. One observer noted during the F2F meeting that participants collaborated effectively, as seen when one individual struggled to find an English word and was quickly assisted by another participant. F2F meetings also had higher levels of spontaneous contribution, where participants were more inclined to share ideas without explicit prompts from the speaker. Unlike VR and MS Teams, where engagement was more passive, F2F meetings created an environment that promoted shared discussion and peer assistance.

5.1.5 Technical issues

Technical issues were only observed in VR, where some participants experienced minor difficulties with movement and seating arrangements. While these issues did not significantly disrupt the meeting, they contributed to an initial period of distraction. No instances of motion sickness were reported, and the audio and video performance remained stable throughout the session. MS Teams and F2F meetings were free of technical issues. The absence of technological barriers in these modalities contributed to a smooth meeting experience, ensuring that discussions flowed without interruptions.

5.1.6 Summary of findings

The qualitative results are summarized in table 5 for a clear overview of the comparison per modality and per theme.

Table 5

Summary of findings, comparison of meeting modalities based on four themes: non-verbal cues, engagement, collaboration, and technical issues.

Theme	VR	MS Teams	F2F
Non-verbal cues	Limited head movement and minimal use of gestures by listeners.	Participants often looked away from the screen, with minimal use of gestures.	Natural use of gestures, nodding, and eye contact.
	The active speaker moved arms a lot, but there was no clear communicative intent for it.	Some nodding was observed, and some participants slouched.	More engagement through facial expressions and gestures.
Engagement	No turn-taking, single speaker.	Some questions were asked, but there was minimal interaction.	Two active speakers and most participants contributed
	Initial distraction due to movement setup.	Some multitasking and inattentiveness were observed.	More laughing and social interactions than VR/MS Teams
Collaboration	The speaker addressed one person by name.	The speaker addressed participants by name.	No name usage, but natural group interaction.
	PowerPoint was the only tool used.	PowerPoint was the only tool used.	PowerPoint was used; participants helped each other.
	No structured turn-taking.	No structured turn-taking.	Smooth turn-taking and peer assistance.
Technical issues	Minor movement/setup issues.	No technical issues.	No technical issues.

5.2 Quantitative analysis

This section presents the quantitative findings of the study on how meeting type (VR, MS Teams, and F2F) impacts participants' perceived social presence and meeting effectiveness. It begins with data preparation and a description of the analytical methods. Next, it presents descriptive statistics, followed by the results of the Repeated Measures ANOVA and linear mixed models. The section concludes with a summary table that presents the key findings for clarity.

5.2.1 Data preparation

Before the analysis, the data were prepared in JASP to ensure accuracy and consistency. Each participant completed the same questionnaire for all three meeting types (VR, MS Teams, and F2F), resulting in repeated measures per individual. The questionnaire responses were first screened for missing values and inconsistencies. Two negatively worded items were reverse-coded to align with the other questions.

A new variable was computed to represent the overall mean score per questionnaire (MeanQ), reflecting participants' general evaluation of social presence of each meeting type. Additional mean scores were calculated for the two outcome variables: degree of perceived social presence (DSP) and effectiveness for the meeting purpose (EMP). To facilitate subtopic-level analysis, average scores for each sub-topic were also computed per meeting type: Co-Presence (COP), Attentional Allocation (ATA), Perceived Message Understanding (PMU), Perceived Affective Understanding (PAU), Perceived Emotional Interdependence (PEI), and Perceived Behavioral Interdependence (PBI).

The data were then reshaped into a wide format for the Repeated Measures ANOVA and linear mixed model analyses in JASP. Participant ID was retained as a grouping variable to account for the within-subject variance. This data preparation ensured that each statistical method was applied to a clean, well-structured dataset.

5.2.2 Method of Analysis

To assess how meeting modality impacts participants' perceived social presence, a within-subjects design was employed, in which each participant rated three different meeting types: VR, MS Teams, and F2F. To analyze the resulting data, three complementary statistical methods were selected to gain a deeper understanding of the effects.

First, Repeated Measures ANOVA was chosen as the primary method to examine whether meeting type had a statistically significant effect on participant responses. This

approach accounts for within-subject variability and is well-suited to experimental designs where the same individuals are exposed to multiple conditions (Dass, 2010).

Second, Subtopic-Level Repeated Measures ANOVA was conducted to explore the effects of meeting type on the six sub-topics. This allowed for an interpretation of which specific sub-scale(s) of social presence was most impacted by meeting modality.

Third, Linear Mixed Models (LMM) were applied as a robustness check to validate the Repeated Measures ANOVA findings while controlling for individual differences in rating tendencies. LMMs are particularly suitable in repeated-measures contexts because they enable flexible modeling of within-subject correlations and can accommodate missing data or heteroscedasticity in responses (Li & Baron, 2012).

5.2.3 Descriptive Results and Measurement Reliability

This section provides an overview of the questionnaire data, presenting descriptive statistics for each meeting modality, evaluating the internal consistency of the measurement scales, and visualizing mean scores along with their corresponding 95% confidence intervals.

Descriptive statistics

Before conducting the statistical tests, descriptive statistics were calculated to gain a preliminary understanding of how participants rated the different meeting modalities. The data indicated a consistent trend across the degree of perceived social presence (DSP), effectiveness for the meeting purpose (EMP), and the mean questionnaire score MeanQ: F2F meetings received the highest ratings, followed by MS Teams, with VR meetings receiving the lowest ratings.

Table 6 presents the descriptive statistics of MeanQ, and Table 7 presents the descriptive statistics for DSP and EMP. These report the number of observations, missing values, means, standard deviations, and minimum and maximum values.

Table 6

Descriptive statistics of the mean questionnaire score (MeanQ) per meeting modality.

	MeanQ		
	VR	*MS Teams*	*F2F*
N	7	7	7
Missing	0	0	0
Mean	2,937	3,522	3,865
Std. Deviation	0,393	0,353	0,478
Minimum	2,278	2,941	3,167
Maximum	3,389	3,889	4,611

Table 7

Descriptive statistics of the grade (1-10) per meeting modality for the degree of social presence (DSP) and the effectiveness for the meeting purpose (EMP).

	DSP			EMP		
	VR	*MS Teams*	*F2F*	*VR*	*MS Teams*	*F2F*
N	7	7	7	7	7	7
Missing	0	0	0	0	0	0
Mean	5,286	7,000	9,429	6,429	8,143	8,857
Std. Deviation	1,890	1,291	1,134	1,988	0,900	0,690
Minimum	2,000	5,000	7,000	4,000	7,000	8,000
Maximum	8,000	9,000	10,000	10,000	10,000	10,000

Internal consistency Cronbach's alpha

To evaluate the internal consistency of the questionnaire's dimensions, Cronbach's alpha (α) was calculated for each subscale. Table 8 presents the reliability coefficients for the six dimensions of the questionnaire. The results indicate that the Co-presence (COP) and Attentional Allocation (ATA) dimensions demonstrated good internal consistency (α = 0,88 and α = 0,86, respectively). The Perceived Emotional Interdependence (PEI) dimension showed acceptable internal consistency (α = 0,71), while Perceived Affective Understanding (PAU) reached borderline acceptable levels (α = 0,65). In contrast, Perceived Message Understanding (PMU) and Perceived Behavioral Interdependence (PBI) showed poor internal consistency (α = 0,45 and α = 0,52, respectively), suggesting that these subscales may not have captured consistent responses across items. These findings should be taken into consideration when interpreting the results of subsequent analyses.

Table 8

Internal consistency (Cronbach's α) of questionnaire dimensions: Co-presence (COP), Attentional Allocation (ATA), Perceived Message Understanding (PMU), Perceived Affective Understanding (PAU), Perceived Emotional Interdependence (PEI), and Perceived Behavioral Interdependence (PBI).

	Cronbach's α	Interpretation
COP	0,88	Good
ATA	0,86	Good
PMU	0,45	Poor
PAU	0,65	Borderline acceptable
PEI	0,71	Acceptable
PBI	0,52	Poor

Note. $\alpha < 0,6$ is considered poor, $0,6 < \alpha < 0,7$ is considered borderline acceptable, $0,7 < \alpha < 0,8$ is considered acceptable, and $\alpha > 0,8$ is considered good.

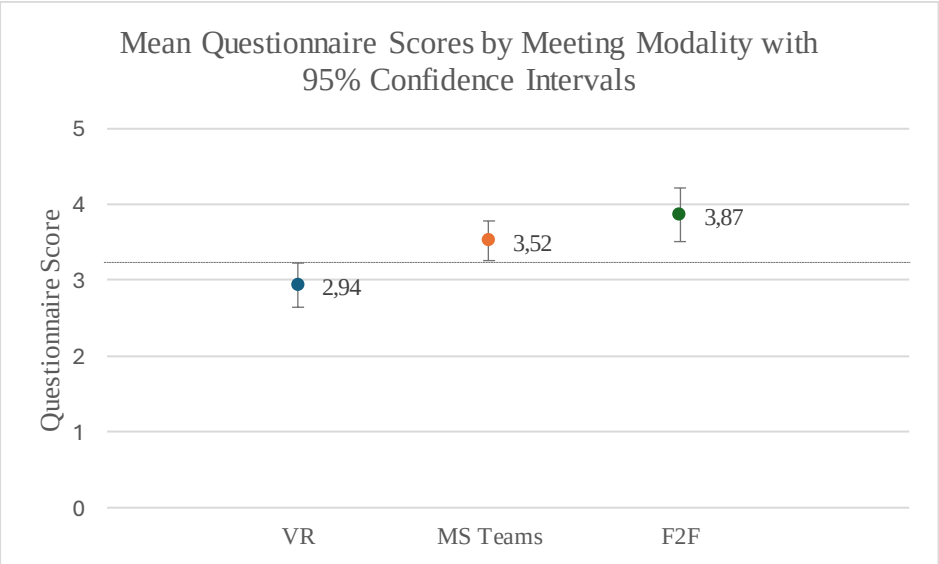
Visualized trend

Figure 8 displays the mean scores for each meeting modality along with 95% confidence intervals to visualize the differences in questionnaire scores across meeting modalities, as shown in the descriptive statistics. This visualization reinforces the pattern observed in the descriptive statistics: VR received the lowest mean rating, followed by MS Teams, and F2F scored the highest. The upper bound confidence interval for VR does not overlap with the lower bounds of MS Teams and F2F.

The results from the descriptive statistics, internal consistency analysis, and confidence intervals provide the basis for the subsequent statistical tests.

Figure 8

Mean questionnaire scores by meeting modality with 95% confidence intervals. VR shows the lowest average score with a confidence interval that does not overlap with those of MS Teams and F2F.



5.2.4 Questionnaire scores Repeated Measures ANOVA

A Repeated Measures ANOVA was conducted for each of the three key outcome variables: the overall mean questionnaire score (MeanQ), the degree of perceived social presence (DSP), and effectiveness for the meeting purpose (EMP). The analysis results are summarized in Table 8, which presents the F-statistic, p-values, effect sizes, applied corrections, and post hoc results for MeanQ, EMP, and DSP. The p-values indicate there was a significant main effect of meeting type for all three outcomes. Mauchly's Test of Sphericity stated a violation of sphericity for EMP, so a Greenhouse-Geisser correction was applied. The corrected degrees of freedom and test statistics are reflected in the table.

Post hoc comparisons revealed that for MeanQ and EMP, VR was rated significantly lower than both MS Teams and F2F, while the difference between MS Teams and F2F was not statistically significant. For DSP, VR was rated significantly lower than F2F but not significantly different from MS Teams. Additionally, MS Teams was rated significantly lower than F2F. The effect size was large for all three variables, according to Cohen (2013). Cohen classifies the effect size as follows: values between 0,01 and 0,059 equal a small effect, values between 0,06 and 0,139 equal a medium effect, and values of 0,14 and above equal a large effect. Table 9 shows that the effect size is above 0,14 for all three variables, indicating that meeting type strongly impacted the scores for each variable. These results are summarized in Table 9, with full ANOVA output and post hoc tables available in Appendix B.

Given this consistent hierarchy in ratings across modalities (VR < MS Teams < F2F), examining which subscales of the questionnaire contributed most to these differences becomes relevant. Therefore, a second set of Repeated Measures ANOVAs was conducted, targeting each subscale individually.

Table 9

Summary of Repeated Measures ANOVA results for the mean questionnaire score (MeanQ), the degree of social presence (DSP), and the effectiveness of the meeting type for the meeting purpose (EMP).

	F (df)	p-value (<0,05)	Effect Size (η^2)	Correction applied	Post hoc results
MeanQ	F(2, 6) = 14,668	<0,001	0,710 (Large)	None	VR<MS Teams, VR<F2F
DSP	F(2, 6) = 21,000	<0,001	0,778 (Large)	None	VR<F2F, MS Teams<F2F
EMP	F(1,068, 6) = 11,355	0,013	0,654 (Large)	Greenhouse-Geisser	VR<MS Teams, VR<F2F

Note. Greenhouse-Geisser correction was applied for EMP due to a violation of sphericity (Mauchly's Test p<0,05).

5.2.5 Sub-topic Repeated Measures ANOVA

A Repeated Measures ANOVA was conducted for each of the six sub-topics: Co-presence (COP), Attentional allocation (ATA), Perceived message understanding (PMU), Perceived affective understanding (PAU), Perceived emotional interdependence, and Perceived behavioral interdependence (PBI). The results of the subtopic-level analysis are summarized in Table 10, which presents the F-statistic, p-values, effect sizes, applied corrections, and post hoc results for each of the six subtopics. The complete set of ANOVA output and post hoc tables is available in Appendix C. The analysis revealed that Co-presence (COP) and Perceived Affective Understanding (PAU) showed significant differences across meeting types, both with large effect sizes ($\eta^2 = 0{,}734$ and $\eta^2 = 0{,}762$, respectively).

For COP, post hoc comparisons revealed that VR was rated significantly lower than both MS Teams and F2F interactions. For PAU, the Greenhouse-Geisser correction was applied due to a violation of sphericity, and results showed that VR was rated significantly lower than both MS Teams and F2F. Additionally, MS Teams was rated significantly lower than F2F.
In contrast, Perceived Emotional Interdependence (PEI) showed no significant difference between the meeting types and had a very small effect size, suggesting this dimension was relatively stable across conditions. Perceived message understanding (PMU) showed a significant main effect ($p = 0{,}028$) with a large effect size ($\eta^2 = 0{,}449$). However, post hoc comparisons did not yield statistically significant pairwise differences. Attentional allocation (ATA) and Perceived behavioral interdependence (PBI) approached significance ($p = 0{,}074$, $p = 0{,}089$, respectively) with large effect sizes ($\eta^2 = 0{,}353$ $\eta^2 = 0{,}332$, respectively), indicating a possible trend.

These results reveal that the largest differences between meeting types occurred in the subtopics of Co-presence (COP) and Perceived Affective Understanding (PAU). Additionally, possible differences were observed in Perceived Message Understanding (PMU), Perceived Behavioral Interdependence (PBI), and Attentional Allocation (ATA), which showed notable effect sizes despite non-significant post hoc comparisons. A linear mixed model analysis was conducted to ensure that these observed differences were not driven by individual variability in participant responses, which is covered in the following section.

Table 10

Summary of Repeated Measures ANOVA results for the subtopics of the questionnaire: Co-presence (COP), Attentional Allocation (ATA), Perceived Message Understanding (PMU), Perceived Affective Understanding (PAU), Perceived Emotional Interdependence (PEI), and Perceived Behavioral Interdependence (PBI).

	F (df)	p-value (<0,05)	Effect Size (η^2)	Correction applied	Post hoc results
COP	F(2, 6) = 16,531	<0,001	0,734 (Large)	None	VR<MS Teams, VR<F2F
ATA	F(2, 6) = 3,252	0,074	0,352 (Large)	None	Not significant
PMU	F(2, 6) = 4,883	0,028	0,449 (Large)	None	No significant post hoc results
PAU	F(1,126, 6) = 19,227	0,003	0,762 (Large)	Greenhouse-Geisser	VR<MS Teams, VR<F2F, MS Teams<F2F
PEI	F(2, 6) = 0,453	0,646	0,070 (Medium)	None	Not significant
PBI	F(2, 6) = 2,984	0,089	0,332 (Large)	None	Not significant

Note. Greenhouse-Geisser correction was applied for EMP due to a violation of sphericity (Mauchly's Test p<0,05).

5.2.6 Linear Mixed Model

The Linear Mixed Models were conducted for the three key outcome variables: overall questionnaire score (MeanQ), the degree of perceived social presence (DSP), and effectiveness for the meeting purpose (EMP). These models included meeting type as a fixed effect and participant as a random intercept, allowing the analysis to adjust for baseline differences in individual rating tendencies.

The results of the LMMs are summarized in Table 11. The full LMMs output is available in Appendix D. All three outcome variables showed a statistically significant main effect of meeting type. The post hoc comparisons indicate that VR was rated significantly lower than both MS Teams and F2F for all three outcome variables. The random intercept variance, which reflects individual variation in baseline scores, was lowest for MeanQ (0,019), indicating that participants' overall scoring tendencies were relatively consistent. In contrast, the variance was higher for DSP (0,786) and EMP (0,730), suggesting that participants differed more in how they graded the degree of social presence and the effectiveness of the meeting purpose across modalities. Despite this variability, the fixed effect of meeting type remained significant in all

cases. JASP used the Satterthwaite approximation to estimate degrees of freedom and test statistics.

Table 11

Summary of the Linear Mixed Models results for the mean questionnaire score (MeanQ), the degree of social presence (DSP), and the effectiveness of the meeting type for the meeting purpose (EMP).

	F (df)	p-value (<0,05)	Significant Differences	Random effects variance
MeanQ	F(2, 12) = 13,253	<0,001	VR<MS Teams VR<F2F	0,116
DSP	F(2, 12) = 11,355	0,002	VR<MS Teams VR<F2F	0,786
EMP	F(2, 12) = 21,000	<0,001	VR<MS Teams VR<F2F	0,730

Note: Fixed effects were tested using the Satterthwaite approximation.

5.2.7 Summary of Findings

Overall, results from the Repeated Measures ANOVA revealed that meeting type significantly impacted participant ratings, with VR receiving lower ratings than MS Teams and F2F. Furthermore, the subtopic-level analysis showed that VR was particularly weak in Co-presence (COP) and Perceived Affective Understanding (PAU). At the same time, no significant differences were found for Attentional Allocation (ATA), Perceived Emotional Interdependence (PEI), Perceived Message Understanding (PMU), and Perceived Behavioral Interdependence (PBI). The linear mixed model confirmed these effects while accounting for individual differences, reinforcing the conclusion that meeting type, rather than personal rating tendencies, explains the observed differences. Table 12 summarizes the key findings from the quantitative analyses to provide a clear overview of the statistical results. The table presents the highest and lowest rated modalities per outcome variable, specifies where statistically significant differences were found, and includes corresponding effect sizes.

Table 12

Summary of findings.

Outcome variable	Highest rated modality	Lowest rated modality	Significant differences	Effect size (η^2)	Interpretation
Mean questionnaire score (MeanQ)	*F2F*	*VR*	*VR<MS Teams*, VR<F2F**	*0,710 (Large)*	*VR rated significantly lower overall*
Degree of perceived social presence (DSP)	*F2F*	*VR*	*VR<F2F* MS Teams<F2F**	*0,778 (Large)*	*F2F>MS Teams>VR*
Effectiveness for meeting purpose (EMP)	*F2F*	*VR*	*VR<MS Teams*, VR<F2F**	*0,654 (Large)*	*VR is perceived as the least effective*
Co-presence (COP)	*F2F*	*VR*	*VR<MS Teams*, VR<F2F**	*0,734 (Large)*	*VR struggles with co-presence*
Attentional Allocation (ATA)	-	-	*Not significant*	*0,352 (Large)*	*No difference across modalities*
Perceived Message Understanding (PMU)	*F2F*	*VR*	*No significant post hoc differences*	*0,449 (Large)*	*Possible difference, not conclusive*
Perceived Affective Understanding (PAU)	*F2F*	*VR*	*VR<MS Teams*, MS Teams<F2F**	*0,762 (Large)*	*F2F is strongest for affective understanding*
Perceived Emotional Interdependence (PEI)	-	-	*Not significant*	*0,070 (Moderate)*	*Perceptions stable across modalities*
Perceived Behavioral Interdependence (PBI)	*F2F*	*VR*	*Not significant*	*0,332 (Large)*	*Possible trend favoring F2F*

Note. indicates a statistically significant difference at p < 0,05. Effect size interpretation: >0,01, <0,06 = small, >0.06, <0,14 = moderate, >0,14 = large.*

6 Analysis

This chapter analyzes the quantitative and qualitative findings of the study by comparing the results of VR, MS Teams, and F2F to prior research on social presence, engagement, and collaboration. Drawing on both the statistical outcomes and the observational data, the analysis aims to determine how the modalities differ, where VR underperforms or aligns with expectations, and how these observations reflect or deviate from existing literature.

6.1 Social Presence Across Modalities

The study found significant differences in perceived social presence between meeting types. The repeated-measures ANOVA revealed that F2F meetings received the highest ratings for social presence, followed by MS Teams, with VR receiving the lowest ratings. Post hoc tests from the mean questionnaire scores (MeanQ) indicated that VR was rated significantly lower than F2F and MS Teams. Post hoc tests on the degree of social presence (DSP) indicated that VR was rated significantly lower than F2F but not significantly lower than MS Teams. However, MS Teams itself was rated significantly lower than F2F, establishing a clear hierarchy: F2F > MS Teams > VR. This pattern contradicts findings in earlier literature, where VR often outperformed traditional video conferencing tools regarding social presence and engagement (Biocca et al., 2003; Van Gent et al., 2024; Hennig-Thurau et al., 2023). The results of this study suggest that, in its current form and setup, VR may lack certain qualities essential for high perceived social presence.

6.2 Subtopic-level insights into social presence

The subtopic-level Repeated Measures ANOVA further revealed the dimensions in which VR meetings fell short. The most significant differences across modalities were observed in co-presence (COP) and perceived affective understanding (PAU), with VR scoring significantly lower than both MS Teams and F2F. In PAU, MS Teams were also rated significantly lower than F2F, further strengthening the observed hierarchy. These findings imply that the immersive nature of VR was insufficient to support the affective and interpersonal components of social presence. This challenges assumptions in previous research (Steinicke et al., 2020), where VR's immersive features were expected to enhance the emotional depth of interactions.

6.3 Explaining the Lower Social Presence Ratings in VR

Several factors likely contributed to VR's underperformance. One key limitation was using static avatars in the Spatial platform, which lack facial expression and eye contact capabilities. These limitations reduce emotional communication and mutual awareness, potentially leading to lower scores in COP and PAU (Singh et al., 2022; Merz et al., 2024). Furthermore, hand tracking within the VR environment was inconsistent. Observers noted that avatar limbs were moving erratically, likely making the avatars feel less realistic. This misalignment may have undermined perceived behavioral interdependence (PBI), as no significant differences were found; however, moderate effect sizes suggested a trend.

6.4 Observed patterns of participant engagement

While F2F meetings demonstrated spontaneous turn-taking and visible engagement (e.g., eye contact, nodding), VR meetings were characterized by passive behavior. The active speaker in VR was not interrupted or supported, while the rest of the group remained quiet and mostly stationary. This is consistent with prior findings suggesting that passive environments hinder collaboration, even in immersive formats (Qiu et al., 2023).

Interestingly, VR meetings had no multitasking compared to MS Teams, where observers noted participants looking away from screens or showing signs of disengagement. However, this lack of distraction in VR did not translate into better engagement or social presence, possibly due to the limited interactivity and social cues.

6.5 Analysis of Meeting Platform Effectiveness

The effectiveness for the meeting purpose (EMP) scores closely mirrored those of the degree of perceived social presence (DSP) and the mean questionnaire scores (MeanQ): F2F was rated most effective, followed by MS Teams and then VR. This result aligns with the notion that the effectiveness of a meeting platform is partially in the quality of social interaction it enables (Hennig-Thurau et al., 2023). Given that most participants were first-time VR users, this may have contributed to lower ratings. Previous research shows that familiarity and training can enhance usability and social presence in VR (Bailenson et al., 2021). The relatively short session durations of only 15 minutes and limited platform exposure likely reduced participants' comfort, limiting the effectiveness of VR.

6.6 Comparison with prior literature

Prior studies generally position VR as a promising tool for enhancing presence beyond what is possible in video conferencing (Biocca et al., 2003; Van Gent et al., 2024; Hennig-Thurau et al., 2023). However, this study challenges that view by showing that VR can underperform under certain conditions, specifically with static avatars, limited gestures, and minimal interactivity. The findings are more aligned with critical perspectives in the literature, which say that technical limitations (avatar realism, tracking fidelity, platform usability) must be addressed before VR can largely replicate F2F or even video conferencing in terms of social presence and effectiveness (Singh et al., 2022; Merz et al., 2024).

The analysis indicates that while VR holds theoretical promise for business communication, its practical implementation in this study fell short. The platform's technical limitations, combined with low familiarity and passive group dynamics, resulted in significantly lower ratings than both MS Teams and F2F across multiple dimensions of social presence. These findings show that immersive technology alone is insufficient; user interaction quality and environment fidelity are just as important for meaningful engagement.

In the next chapters, the key findings from this analysis are discussed and translated into practical implications and recommendations for businesses considering the adoption of VR for meetings. This is followed by a discussion of the study's limitations.

7 Discussion

This study examined how VR, MS Teams, and F2F differ in perceived social presence and effectiveness in a professional context. As detailed in the previous chapter, the results consistently showed that VR underperformed across all measurements. The 95% confidence intervals in Figure 8 further validated these results, demonstrating the robustness of the study. Despite a small sample size, these outcomes indicate that VR is less effective than MS Teams and F2F with the current meeting setup.

The subtopic analysis showed that VR scored lower on Co-presence (COP) and Perceived Affective Understanding (PAU). These subtopics are closely tied to emotional connection and interpersonal awareness. To provide a broader perspective, the immersion of VR does not inherently offer good simulated in-person social and emotional interactions. Due to VR's limitations in this study, the 3D environments prevented users from feeling present with one another. This finding carries practical implications for organizations considering VR as a substitute for physical meetings. Based on this small sample size, VR seems not to be the most effective tool for maintaining the social and emotional aspects of professional interactions.

Furthermore, the observers noted that the VR meeting was mainly one-sided; a single speaker dominated it, and no turn-taking or non-verbal cues were observed. This could be due to the meeting setup, which was giving a presentation on a work-related topic, not as interactive as a brainstorming session, for example. However, the F2F session also involved a presentation on a work-related topic. During the F2F meeting, observers noted much interaction between participants, use of gestures, eye contact, and natural conversation flow. This shows that the meeting format alone cannot explain the one-sided nature of the VR meeting.

This study argues that the underperformance of VR in providing perceived social presence stems from the technological and experiential limitations of the VR platform. In its current state, the VR platform allows only for limited ways to convey messages. The avatars used in the Spatial platform did not allow for facial expressions and eye-tracking, causing the faces of the avatars to be static. Furthermore, the avatars did not have accurate hand-tracking, causing erratic hand movements during the VR meeting. These issues were compounded by the hardware constraints of the Meta Quest 2 headset. These two limitations made it difficult for participants to convey messages through non-verbal communication, limiting the communicative effectiveness.

VR does make it hard for participants to multitask due to its immersive nature, which reduces distractions. However, it did not compensate for the lack of nonverbal communication, essential for effective group interaction. This finding aligns with the literature that social presence is not only about being immersed in the same digital environment but also about effectively conveying social cues, such as nodding, eye contact, and body language. MS Teams also outperformed VR in this study despite being less immersive. MS Teams shows the real faces of participants instead of the static avatar faces, allowing for conveying messages through facial expressions. This indicates that for this group immersion alone is insufficient for effective group interaction. Instead, a balance between technical fidelity and conveying messages is required for a virtual environment to be effective and have a high social presence.

Exploring this further, VR's limitations also indicated that participants' perceived behavioral interdependence was low in VR. Behavioral interdependence is a subcomponent of social presence, referring to participants' influence on each other's behavior. This was low due to the avatars' lack of expressiveness. Participants could not effectively convey messages, let alone effectively receive them, making it evident that they could not influence each other's behavior. This further reinforces that social presence in a virtual environment requires the ability to verbally and nonverbally convey messages.

Next to the technical limitations of VR, user experience could have impacted the results. The participants in this study only received a short 30-minute training to get accustomed to the environment and the controls. Most participants also had no experience in VR, making this their first time. Additionally, the reliability of the measurement instrument adds nuance to the interpretation of the results. Several subtopics showed low internal consistency, as reflected by their Cronbach's alpha scores. This may have been caused by the shortened version of the original questionnaire used in this study. Fewer questionnaire items reduce the ability of the questions to capture each construct fully. Furthermore, the original questionnaire showed good Cronbach's alpha scores for each construct. However, the lack of back-and-forth communication in VR and MS Teams may have also led to the lower reliability scores. The questions from the subtopics with low reliability required interactivity and engagement, which were not there. Therefore, participants likely answered these questions similarly with little response variability, reducing the internal reliability of these subtopics.

The literature views VR as the future of remote collaboration; however, this study's findings caution against assuming that more immersive technologies will automatically translate into better communication. Social presence is not just about being virtually together; it is about being mutually understood.

8 Conclusion

This study examined how VR, MS Teams, and F2F differ in perceived social presence and effectiveness within a professional context. Participants participated in an experiment in their business environment, experiencing all three meeting platforms. The adapted version of the Networked Minds Social Presence Inventory questionnaire served as the measurement tool, complemented by observations. The results revealed a consistent pattern: F2F received the highest ratings, followed by MS Teams, with VR rated the lowest across all measurements.

The findings showed that while VR offers an immersive experience and reduces distractions during meetings, immersion alone is insufficient to deliver a high perceived social presence and effective group interaction. The observational data confirmed that participants had difficulty conveying messages through nonverbal communication. The subtopic-level analysis also showed low scores on co-presence and perceived affective understanding, meaning there was a lack of emotional and interpersonal connection in VR. Importantly, these findings challenge the assumption that more immersion automatically leads to better communication. VR offers great immersion by simulating 3D environments. However, as this study showed, VR cannot convey nonverbal messages through gestures and facial expressions in the current setup. These limitations and the participants' inexperience with VR contributed to VR's underperformance in this study.

The study also reflects on the measurement tools and small sample size. Some subtopics had strong internal consistency; others also showed low Cronbach's alpha scores. The sample size consisted of only seven participants, which is relatively small. However, despite the small sample size, the study showed consistent results. The non-overlapping confidence intervals supported these results, significant repeated measures ANOVA findings, and large effect sizes. These results led to meaningful and actionable insights. Social presence is not achieved simply by co-locating people in a virtual space but by enabling mutual understanding through clear and effective verbal and nonverbal communication.

From a practical standpoint, these results suggest that organizations should be cautious about adopting VR as a substitute for video conferencing. VR platforms must evolve, adding the ability to convey nonverbal messages through accurate gestures and facial expressions. A good place to start is enabling eye-tracking because the eyes tell a lot about the intentions and thoughts of human beings in communication. Users of VR also need to become more experienced in navigating the environments. Until then, VR should be seen as a complementary tool, not a full replacement for video conferencing.

In conclusion, this study contributes to the growing body of knowledge on social presence and interaction. This study has shown that in the current meeting setup, it is not the level of immersion that defines social presence but the mutual understanding created by effectively conveying verbal and nonverbal messages.

8.1 Practical implications

For businesses considering the adoption of VR for their professional meetings, the following practical implications from the findings of this study are suggested.

First, businesses should invest in the latest VR devices because the quality of social interactions in VR is heavily influenced by the realism of the VR environment and the avatars. For effective meetings, the VR headsets should have facial and eye-tracking capabilities, and proper hand-tracking. The VR environment should also have high-quality textures and proper lighting.

Second, to get the most out of VR, businesses should take full advantage of the interactive capabilities of VR. This includes utilizing virtual whiteboards, object manipulations, and collaborative environments instead of solely relying on a PowerPoint presentation for the meetings. These tools can transform VR meetings from passive viewing experiences into active, engaging sessions.

Third, employees should receive proper training before VR can be utilized to its full potential in business meetings. The training should teach employees how to use all the interactive capabilities of VR. Just like they learned how to screenshare and create break-out rooms in MS Teams.

Lastly, while the VR headsets help reduce distractions and multitasking by their immersive nature, businesses must ensure that the meetings are designed to encourage active participation. Simply placing employees in a VR environment does not guarantee engagement. The meetings should be structured to encourage dialogue, interaction, and collaboration to take advantage of VR.

8.2 Limitations and Future Research

This study provides valuable insights, however limitations should be considered, which can be accounted for in future research.

- Sample Size: the study was conducted with a small sample (n=7), limiting the generalizability of findings. Future research should include a larger sample size to improve statistical power. The subtopic Repeated Measures ANOVA showed that for

Perceived Message Understanding (PMU) and Perceived Behavioral Interdependence (PBI), there was a large effect size but no significant post hoc result. With a larger sample size, these effects might have been significant.

- Short meeting duration: each meeting lasted only 15 minutes. Longer VR sessions might yield different results because participants might need time to get used to the environment.
- The meetings were structured around a PowerPoint presentation with one or two active speakers. This experiment format may not reflect the full scope of real-world business meetings, such as problem-solving, brainstorming, or decision-making tasks. Future research should incorporate more diverse and interactive meeting formats to test how well VR supports these types of meeting formats.
- Platform constraints: the study was conducted on Spatial, a VR platform that lacks facial tracking and good gesture recognition. Future studies should compare multiple VR platforms to determine whether more advanced systems could improve social presence.
- This study used Meta Quest 2 devices which are outdated. These devices had hand-tracking issues during the experiment taking away from the realness of the experience for the participants. Future research should use the latest VR devices. Devices that allow for eye-tracking and proper hand-tracking.
- First-time users: most participants had limited VR experience, which could have influenced their comfort levels and engagement. Studies involving experienced VR users may yield different results because experienced users can use the interactive tools available in VR.

9 Reference list

This reference list consists of an academic reference list and an Industry, News, and Business reference list

9.1.1 Academic Reference List

Ahn, S. J., Levy, L., Eden, A., Won, A. S., MacIntyre, B., & Johnsen, K. (2021). IEEEVR2020: Exploring the first steps toward standalone virtual conferences. *Frontiers in Virtual Reality*, *2*, 648575. Doi: 10.3389/frvir.2021.648575

Baloian, N., Geisse, F., & Rubilar, J. (2023). The mind in virtual meetings: Comparing VR and video conferencing using EEG. *Journal of Cognitive Presence, 28*(2), 141-158. Doi: 10.1007/978-3-031-48306-6_26

Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments, 12*(5), 456-480. Doi: 10.1162/105474603322761270

Chessa, M., & Solari, F. (2021). The sense of being there during online classes: A usability analysis of web-conferencing and virtual reality platforms. *Educational Presence Studies, 7*(3), 203-220. Doi: 10.1080/0144929X.2021.1957017

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge. Doi: 10.4324/9780203771587

Cummings, J. J., & Bailenson, J. N. (2016). How immersive is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology, 19*(2), 272-309. Doi: 10.1080/15213269.2015.1015740

Dass, M. (2010). Repeated Measures ANOVA. *Wiley International Encyclopedia of Marketing*. Doi: 10.1002/9781444316568.WIEM02070.

Dean, T., Green, H., & Walters, S. (2014). Refining personal and social presence in virtual meetings. *International Journal of Social Computing, 8*(1), 15-27. Doi: 10289/9331

Harms, C., & Biocca, F. (2004). Internal consistency and reliability of the networked minds measure of social presence.

Hennig-Thurau, T., et al. (2023). Social interactions in the metaverse: Framework, initial evidence, and research roadmap. *Journal of Digital Interaction Studies, 42*(1), 17-33. Doi: 10.1007/s11747-022-00908-0

Higgins, D., Fribourg, R., & McDonnell, R. (2021). Remotely perceived: Investigating the influence of valence on self-perception and social experience for dyadic video-conferencing with personalized avatars. Frontiers in Virtual Reality, 2, 668499. Doi: 10.3389/frvir.2021.668499

Kimmel, S., Landwehr, E., & Heuten, W. (2024). Kinetic Connections: Exploring the Impact of Realistic Body Movements on Social Presence in Collaborative Virtual Reality. Proceedings of the ACM on Human-Computer Interaction, 8(CSCW2), 1-30. Doi: 10.1145/3686910

Lawrence, E., et al. (2021). Project Starline: A high-fidelity telepresence system. *Telepresence Engineering Review, 18*(2), 98-116. Doi: 10.1145/3641517.3664381

Li, Y., Baron, J., Li, Y., & Baron, J. (2012). Linear mixed-effects models in analyzing repeated-measures data. *Behavioral research data analysis with R*, 177-204. Doi: 10.1007/978-1-4614-1238-0_10

Merz, R., et al. (2024). Universal access for social XR across devices: The impact of immersion on asymmetric collaboration. *Journal of Extended Realities, 12*(1), 50-73. Doi: 10.1109/VR58804.2024.00105

Qiu, J., et al. (2023). ViGather: Inclusive virtual conferencing across traditional screen devices and mixed reality. *Journal of Hybrid Communication, 19*(2), 88-102. Doi: 10.1145/3604279

Rzayev, A., et al. (2020). It's not always better when we're together: Effects of being accompanied in virtual reality. *Journal of Experimental Virtual Studies, 10*(3), 175-189. Doi: 10.1145/3334480.3382826

Short, J., Williams, E., & Christie, B. (1976). The social psychology of telecommunications. London: John Wiley & Sons.

Singh, R., et al. (2022). Engagement and quality of experience in remote business meetings: A social VR study. *Business Presence Research, 8*(4), 250-266. Doi: 10.1145/3552483.3556457

Speidel, R., et al. (2023). Virtual reality against Zoom fatigue: Field study on teaching and learning experience. *Educational Technology Quarterly, 39*(1), 100-118. Doi: 10.3205/zma001601

Steinicke, F., et al. (2020). A first pilot study to compare virtual group meetings using video conferences and immersive virtual reality. *Journal of Virtual Meeting Sciences, 14*(2), 23-38. Doi: 10.1145/3385959.3422699

Wang, F., et al. (2024). Socially late, virtually present: Transforming asynchronous social interactions in VR. *Journal of Social XR Studies, 31*(1), 67-81. Doi: 10.1145/3613904.3642244

9.1.2 Industry, News, and Business Reference List

Apple. (n.d.). *Apple Vision Pro*. Retrieved from https://www.apple.com/apple-vision-pro/

Eurofound. (2022). *Living, working and COVID-19 dataset – 2022 edition*. Publications Office of the European Union. Retrieved from: https://www.eurofound.europa.eu/data/living-working-covid-19

Google. (n.d.). *Project Starline*. Retrieved from: https://blog.google/technology/research/project-starline/

McKinsey & Company. (2022). *Americans are embracing flexible work—and they want more of it*. Retrieved from: https://www.mckinsey.com/industries/real-estate/our-insights/americans-are-embracing-flexible-work-and-they-want-more-of-it

Meta. (n.d.). *Horizon Worlds*. Retrieved from https://horizon.meta.com/

Meta. (n.d.). *Reality Labs: Building the future of connection in the metaverse*. Retrieved from https://about.meta.com/realitylabs/

Microsoft. (n.d.). *HoloLens: Mixed reality for enterprise*. Retrieved from https://www.microsoft.com/nl-NL/hololens?msockid=1e02480a4f84664d36d55b8f4e106768

Microsoft. (n.d.). *Microsoft Mesh: Enabling collaboration in mixed reality*. Retrieved from https://www.microsoft.com/en-us/microsoft-teams/microsoft-mesh?msockid=1e02480a4f84664d36d55b8f4e106768

NVIDIA. (n.d.). *Omniverse: A platform for real-time 3D collaboration and simulation*. Retrieved from https://www.nvidia.com/en-us/omniverse/

Reuters. (2024, May 30). *Google and Magic Leap partner for immersive AR experiences*. Retrieved from https://www.reuters.com/technology/google-augmented-reality-startup-magic-leap-strike-partnership-deal-2024-05-30/

The Verge. (2022, January 20). *Google's Project Iris AR headset aimed for 2024*. Retrieved from https://www.theverge.com/2022/1/20/22892152/google-project-iris-ar-headset-2024

The Verge. (2024, November 30). *Fortnite Remix: The Finale concert featuring Juice WRLD, Snoop Dogg, and more*. Retrieved from

https://www.theverge.com/2024/11/30/24308372/fortnite-remix-finale-concert-juice-wrld-chapter-6-how-to-watch

The Verge. (2024, October 30). *Google Play Store preps for AR and VR headset integration*. Retrieved from https://www.theverge.com/2024/10/30/24283894/google-play-store-ar-vr-xr-headset-smart-glasses-support-code-leak

The Verge. (2024, October 5). *Epic Games and Unreal Engine: Paving the way for metaverse innovation*. Retrieved from https://www.theverge.com/2024/10/5/24262376/epic-unreal-engine-6-fortnite-metaverse-plans

The Verge. (2024, September 5). *Qualcomm, Samsung, and Google collaborate on mixed reality glasses*. Retrieved from https://www.theverge.com/2024/9/5/24236545/qualcomm-mixed-reality-smart-glasses-google-samsung

Unity Technologies. (n.d.). *The road to the metaverse: Unity's impact on 3D content creation*. Retrieved from https://create.unity.com/road-to-metaverse

10 Appendix

10.1 Appendix A

10.1.1 Industry leaders in VR and-metaverse development

The development of VR and the metaverse is being shaped by several leading technology companies that are investing a lot of money and using their expertise to create immersive virtual environments and tools. We will now provide examples of what the industry leaders are doing in the field of the metaverse and VR.

Meta Platforms (previously Facebook) established itself as a leader in the next generation of social VR and metaverse. Meta supports interaction, content creation and shared experiences in virtual spaces through its Horizon Worlds platform[1]. Its Reality Labs segment also works on developing VR and AR technologies, like the consumer- and enterprise-grade Meta Quest line of VR headsets[2]. They've just launched the Quest 3S, which is effectively a cheaper version of the Quest 3, making it clear that meta wants VR to become more accessible to consumers.

While Meta Platforms brought social interaction to the metaverse, NVIDIA used its knowledge of high-performance computing and graphics to make metaverse contributions. Its Omniverse platform offers an open, real-time simulation and collaboration environment for professionals in fields like manufacturing, architecture and entertainment. By integrating tools for 3D design and visualization with its cutting-edge GPUs, NVIDIA enhances the realism and interactivity of virtual experiences. The company's innovations in real-time ray tracing and AI-powered virtual assistants further position it as a key player in metaverse development[3].

Microsoft is also making significant investments in VR and AR through its Mesh platform and HoloLens devices. Microsoft Mesh enables collaboration in shared virtual spaces, by combining holographic experiences with traditional VR and AR[4]. This platform integrates seamlessly with MS Teams, allowing businesses to collaborate in 3D. The HoloLens is a standalone holographic device with enterprise-ready applications to enhance user accuracy and output. It offers advanced features such as spatial mapping and gesture control, making it a tool for enterprise applications, including training and design[5].

---

[1] Retrieved from: Meta Horizon
[2] Retrieved from: Reality Labs | Meta
[3] Retrieved from: Omniverse Platform for OpenUSD | NVIDIA
[4] Retrieved from: Introducing Microsoft Mesh | Connect like never before
[5] Retrieved from: Microsoft HoloLens | Mixed reality-technologie voor bedrijven

Apple recently entered the VR market with its Apple Vision Pro. This spatial computing device merges VR and AR to support remote collaboration, media creation, and content consumption. Its features, such as eye-tracking and hand-gesture recognition, provide intuitive user interactions that align with Apple's commitment to creating seamless and accessible technologies. The Vision Pro highlights Apple's vision of using immersive technologies to enhance productivity and creativity[6].

Google has introduced Project Starline[7], a groundbreaking video communication system that uses 3D modelling for natural interaction without VR headsets. Google has also partnered with Magic Leap[8] to develop immersive AR experiences and is collaborating with Samsung and Qualcomm[9] on mixed reality smart glasses. Furthermore, google is integrating AR and VR support into the Google Play Store, expanding its ecosystem for immersive applications[10].

Epic Games has used its gaming expertise to contribute to the metaverse. The company has hosted virtual concerts and events within the game Fortnite, showing the potential for immersive entertainment experiences. For example, the "Remix: The Finale" event featured artists like Snoop Dogg, Eminem, Ice Spice, and Juice WRLD[11]. Furthermore, its Unreal Engine has become a great tool for developers creating interactive 3D environments across industries[12].

Finally, Unity Technologies is a leading provider of real-time 3D development platforms. They enable developers to build metaverse applications[13]. The Unity Engine, a widely used tool for creating interactive 3D content, supports industries ranging from gaming to education.

---

[6] Retrieved from: Apple Vision Pro - Apple
[7] Retrieved from: Project Starline: Feel like you're there, together
[8] Retrieved from: Exclusive: Google, augmented reality startup Magic Leap strike partnership deal | Reuters
[9] Retrieved from: Qualcomm's mixed reality project with Samsung and Google is... glasses - The Verge
[10] Retrieved from: Google doesn't sell headsets anymore — but its app store is getting ready for them - The Verge
[11] Retrieved from: Fortnite Remix: The Finale concert: how to watch - The Verge
[12] Retrieved from: Epic has a plan for the rest of the decade - The Verge
[13] Retrieved from: Road to the Metaverse | Unity

10.2 Appendix B

The Repeated Measures ANOVA revealed a significant main effect of meeting type on participant ratings, $F(2,6) = 10,498$, $p = 0,002$, shown in table 13. Post hoc comparisons using Bonferroni correction indicated that VR meetings were rated significantly lower than both MS Teams and F2F ($p < 0,05$), while no significant differences were found between MS Teams and F2F ($p > 0,05$), shown in table 14. Descriptive statistics showed that F2F meetings had the highest ratings ($M = 3,746$), followed by MS Teams ($M = 3,476$), with VR receiving the lowest ratings ($M = 2,937$). A similar pattern was found for Social Presence (DSP) and Effectiveness (EMP), where VR was rated significantly lower than MS Teams and F2F. The tables for the DSP and EMP are shown in Appendix B.

Cases	Sum of Squares	df	Mean Square	F	p	η^2
Meeting type	2.379	2	1.189	10.498	0.002	0.636
Residuals	1.359	12	0.113			

Table 13: outcome Repeated Measures ANOVA for the meeting type

The effect size for the Repeated Measures ANOVA was $\eta^2 = 0,636$, indicating a large effect according to Cohen's guidelines. This suggests that meeting type had a strong impact on participants ratings, with VR consistently receiving lower ratings than the other two modalities.

		Mean Difference	SE	df	t	p_{bonf}
VR	Teams	−0.540	0.132	6	−4.078	0.020
	F2F	−0.810	0.215	6	−3.769	0.028
Teams	F2F	−0.270	0.183	6	−1.475	0.572

Table 14: post hoc comparisons for the meeting type

A boxplot was created to visually represent the differences in ratings across meeting types. The boxplot, depicted in figure 9, illustrates the distribution, median, and variability in scores for VR, MS Teams, and F2F meetings. These results suggest that participants prefer F2F and MS Teams meetings over VR.

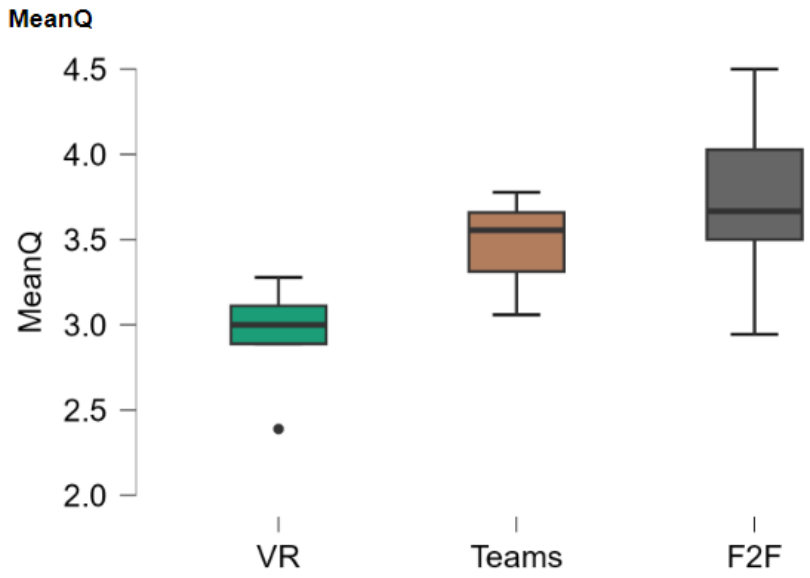


Figure 9: boxplot of the mean questionnaire score per meeting modality.

Table 15-18 depict the results of the Repeated Measures ANOVA for DSP and EMP. Figure 10 depicts the boxplot of DSP (left) and EMP (right).

Cases	Sum of Squares	df	Mean Square	F	p	η^2
DSP	60.667	2	30.333	21.000	< .001	0.778
Residuals	17.333	12	1.444			

Table 15: Repeated Measures ANOVA results for the perceived social presence.

Post Hoc Comparisons - DSP

		Mean Difference	SE	df	t	p_{bonf}
DSP_VR	DSP_Teams	−1.714	0.606	6	−2.828	0.090
	DSP_F2F	−4.143	0.705	6	−5.879	0.003
DSP_Teams	DSP_F2F	−2.429	0.612	6	−3.970	0.022

Table 16: post hoc comparisons for the perceived social presence.

Cases	Sum of Squares	df	Mean Square	F	p	η^2
EMP	21.810[a]	2[a]	10.905[a]	11.355[a]	0.002[a]	0.654
Residuals	11.524	12	0.960			

Note. Type III Sum of Squares

[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated (p < .05).

Table 17: Repeated Measures ANOVA results for the effectiveness for the meeting purpose.

		Mean Difference	SE	df	t	p_{bonf}
EMP_VR	EMP_Teams	−1.714	0.474	6	−3.618	0.033
	EMP_F2F	−2.429	0.719	6	−3.378	0.045
EMP_Teams	EMP_F2F	−0.714	0.286	6	−2.500	0.140

Table 18: post hoc comparisons for the effectiveness for the meeting purpose.

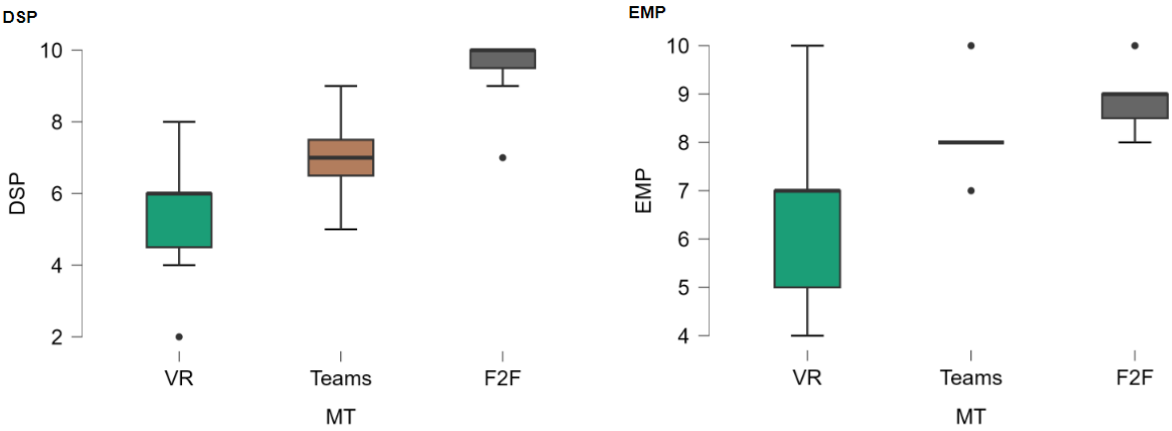


Figure 10: boxplot DSP (left) and EMP (right).

10.3 Appendix C

This appendix provides a detailed breakdown of the Repeated Measures ANOVA results for each subtopic of the questionnaire. The analysis examines whether meeting type (VR, MS Teams, and F2F) significantly influenced participant responses in the following six categories:

- Co-Presence (COP)
- Attentional Allocation (ATA)
- Perceived Message Understanding (PMU)
- Perceived Affective Understanding (PAU)
- Perceived Emotional Interdependence (PEI)
- Perceived Behavioral Interdependence (PBI)

For each subtopic, we report the F-statistic, effect size (η^2), and post hoc comparisons where applicable.

10.3.1 Repeated Measures ANOVA for subtopic-level analysis

To identify which specific aspects of the meeting experience were most affected by meeting type, a Repeated Measures ANOVA was conducted for each subtopic of the questionnaire separately: COP, ATA, PMU, PAU, PEI, and PBI. This approach helps identify which meeting aspects contribute most to the overall differences observed in previous analyses.

Post hoc comparisons revealed that VR was rated significantly lower than both MS Teams and F2F for co-presence ($p = 0{,}018$) and perceived affective understanding ($p = 0{,}009$). In contrast, no significant differences were found for attentional allocation, perceived emotional interdependence, or perceived behavioral interdependence. These findings suggest that VR meetings struggle primarily in maintaining co-presence and affective understanding, while attentional focus and emotional engagement remain comparable across all meeting types.

10.3.2 Repeated Measures ANOVA COP

The Repeated Measures ANOVA for Co-Presence (COP) revealed a significant main effect of meeting type, $F(2,12) = 16,531$, $p < ,001$, $\eta^2 = 0.,734$, indicating a large effect. (Table 19).

Post Hoc comparisons (Table 20):

VR < MS Teams (p = 0,018)

VR < F2F (p = 0,007)

MS Teams vs. F2F (p = 0,488, not significant)

Interpretation: VR significantly reduces the feeling of co-presence compared to MS Teams and F2F meetings. F2F meetings had the highest co-presence ratings, followed by MS Teams, with VR scoring the lowest.

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
COP	10.649	2	5.324	16.531	< .001	0.734
Residuals	3.865	12	0.322			

Note. Type III Sum of Squares

Table 19: Repeated Measures ANOVA COP result.

Descriptives

Descriptives

COP	N	Mean	SD	SE	Coefficient of variation
VR	7	3.321	0.886	0.335	0.267
MS Teams	7	4.571	0.713	0.269	0.156
F2F	7	5.000	0.000	0.000	0.000

Table 20: Descriptive statistics COP

Post Hoc Tests

Post Hoc Comparisons - COP

		Mean Difference	SE	df	t	p_{bonf}
VR	MS Teams	-1.250	0.302	6	-4.138	0.018
	F2F	-1.679	0.335	6	-5.010	0.007
MS Teams	F2F	-0.429	0.269	6	-1.591	0.488

Table 21: Post Hoc tests results COP.

10.3.3 Repeated Measures ANOVA ATA

The Repeated Measures ANOVA for attentional allocation (ATA) showed no significant effect of meeting type, F(2,12) = 0,026, p = ,974, η^2 = 0,004, indicating a negligible effect. (Table 22).

Interpretation: participants allocated their attention equally across all three meeting types, suggesting that VR does not inherently reduce attentional allocation compared to MS Teams or F2F meetings.

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
ATA	0.011	2	0.005	0.026	0.974	0.004
Residuals	2.434	12	0.203			

Note. Type III Sum of Squares

Table 22: Repeated Measures ANOVA results ATA.

Descriptives

Descriptives

ATA	N	Mean	SD	SE	Coefficient of variation
VR	7	3.429	0.600	0.227	0.175
MS Teams	7	3.429	0.460	0.174	0.134
F2F	7	3.381	0.525	0.198	0.155

Table 23: Descriptive statistics ATA.

10.3.4 Repeated Measures ANOVA PMU

The Repeated Measures ANOVA for perceived message understanding (PMU) showed a significant effect of meeting type, $F(2,12) = 4,883$, $p = ,028$, $\eta^2 = 0,449$, indicating a moderate effect. (Table 24). However, post hoc comparisons did not show significant differences between specific meeting types.

Post Hoc comparisons (Table 25):

VR vs. MS Teams ($p = 0,155$, not significant)

VR vs. F2F ($p = 0,155$, not significant)

MS Teams vs. F2F ($p = 0,283$, not significant)

Interpretation: while F2F and MS Teams had higher PMU scores than VR, the differences were not statistically significant, suggesting that VR does not drastically impair perceived message understanding.

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
PMU	2.127	2	1.063	4.883	0.028	0.449
Residuals	2.614	12	0.218			

Note. Type III Sum of Squares

Table 24: Repeated Measures ANOVA results PMU.

Descriptives

Descriptives

PMU	N	Mean	SD	SE	Coefficient of variation
VR	7	3.905	0.460	0.174	0.118
MS Teams	7	4.429	0.568	0.215	0.128
F2F	7	4.667	0.509	0.192	0.109

Table 25: Descriptive statistics PMU.

Post Hoc Tests

Post Hoc Comparisons - PMU

		Mean Difference	SE	df	t	p_{holm}
VR	MS Teams	-0.524	0.216	6	-2.420	0.155
	F2F	-0.762	0.315	6	-2.421	0.155
MS Teams	F2F	-0.238	0.202	6	-1.179	0.283

Note. P-value adjusted for comparing a family of 3 estimates.

Table 26: Post Hoc tests results PMU.

10.3.5 Repeated Measures ANOVA PAU

The Repeated Measures ANOVA for perceived affective understanding (PAU) revealed a significant effect of meeting type, $F(2,12) = 19,227$, $p = 0,003$, $\eta^2 = 0,762$, indicating a large effect. (Table 27).

Post Hoc comparisons (Table 29):

VR < MS Teams ($p = 0,011$)

VR < F2F ($p = 0,009$)

MS Teams < F2F ($p = 0,011$)

Interpretation: VR significantly reduced participants' ability to understand affective cues compared to MS Teams and F2F meetings. F2F meetings facilitated the best affective understanding, followed by MS Teams, with VR scoring the lowest.

Within Subjects Effects

Cases	Sphericity Correction	Sum of Squares	df	Mean Square	F	p	η^2
PAU	Greenhouse-Geisser	18.193	1.126	16.154	19.227	0.003	0.762
Residuals	Greenhouse-Geisser	5.677	6.757	0.840			

Note. Type III Sum of Squares

[a] Mauchly's test of sphericity indicates that the assumption of sphericity is violated ($p < .05$).

Table 27: Repeated Measures ANOVA results PAU.

Descriptives

Descriptives

PAU	N	Mean	SD	SE	Coefficient of variation
VR	7	1.952	0.678	0.256	0.348
MS Teams	7	3.595	0.652	0.246	0.181
F2F	7	4.143	0.742	0.280	0.179

Table 28: Descriptive statistics PAU.

Post Hoc Tests

Post Hoc Comparisons - PAU

		Mean Difference	SE	df	t	p_{holm}
VR	MS Teams	-1.643	0.428	6	-3.839	0.011
	F2F	-2.190	0.453	6	-4.831	0.009
MS Teams	F2F	-0.548	0.130	6	-4.223	0.011

Note. P-value adjusted for comparing a family of 3 estimates.

Table 29: Post Hoc tests results PAU.

10.3.6 Repeated Measures ANOVA PEI

The Repeated Measures ANOVA for perceived emotional interdependence (PEI) showed no significant effect of meeting type, $F(2,12) = 0,453$, $p = ,646$, $\eta^2 = 0,070$, indicating a small effect. (Table 30).

Interpretation: meeting type did not significantly impact participants' sense of emotional interdependence, suggesting that VR, MS Teams, and F2F meetings provide a similar level of emotional connection.

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
PEI	1.153	2	0.577	0.453	0.646	0.070
Residuals	15.291	12	1.274			

Note. Type III Sum of Squares

Table 30: Repeated Measures ANOVA results PEI.

Descriptives

Descriptives

PEI	N	Mean	SD	SE	Coefficient of variation
VR	7	2.190	0.997	0.377	0.455
MS Teams	7	1.952	0.826	0.312	0.423
F2F	7	2.524	1.609	0.608	0.637

Table 31: Descriptive statistics PEI.

10.3.7 Repeated Measures ANOVA PBI

The Repeated Measures ANOVA for perceived behavioral interdependence (PBI) showed no significant effect of meeting type, $F(2,12) = 2,984$, $p = ,089$, $\eta^2 = 0,332$, indicating a moderate effect but no significant post hoc differences. (Table 34)

Post Hoc comparisons (Table 33):

VR vs. MS Teams (p = 0,105, not significant)

VR vs. F2F (p = 0,209, not significant)

MS Teams vs. F2F (p = 0,751, not significant)

Interpretation: there was a trend suggesting differences, but these differences were not statistically significant. This suggests that meeting type may influence behavioral interdependence.

Within Subjects Effects

Cases	Sum of Squares	df	Mean Square	F	p	η^2
PBI	2.952	2	1.476	2.984	0.089	0.332
Residuals	5.937	12	0.495			

Note. Type III Sum of Squares

Table 32: Repeated Measures ANOVA results PBI.

Descriptives

Descriptives

PBI	N	Mean	SD	SE	Coefficient of variation
VR	7	2.048	0.705	0.267	0.344
MS Teams	7	2.905	0.763	0.288	0.263
F2F	7	2.762	0.810	0.306	0.293

Table 33: Descriptive statistics PBI.

Post Hoc Tests

Post Hoc Comparisons - PBI

		Mean Difference	SE	df	t	p_{holm}
VR	MS Teams	-0.857	0.316	6	-2.714	0.105
	F2F	-0.714	0.374	6	-1.910	0.209
MS Teams	F2F	0.143	0.429	6	0.333	0.751

Note. P-value adjusted for comparing a family of 3 estimates.

Table 34: Post Hoc tests results PBI.

10.4 Appendix D

Table 34-39 depict the results of the linear mixed model for DSP and EMP.

10.4.1 Linear mixed model

A Linear Mixed Model (LMM) was conducted to analyze the effect of meeting type on participant ratings while accounting for individual differences. This was also done the same way for DSP and EMP, of which the tables and figures are shown in Appendix D. The model included meeting type as a fixed effect and each participant as a random intercept to control for individual rating tendencies. Results showed a significant main effect of meeting type on scores, F(2,12) = 9,435, p = 0,003, shown in table 35.

Effect	df	F	p
MT	2, 12.00	9.435	0.003

Note. Model terms tested with Satterthwaite testMethod.

Table 35: result from the model.

Fixed Effects Estimates

Term	Estimate	SE	df	t	p
Intercept	3.389	0.092	6.000	36.936	< .001
MT (1)	−0.437	0.106	12.000	−4.105	0.001
MT (2)	0.087	0.106	12.000	0.822	0.427

Table 36: fixed effect estimates for the meeting type.

Fixed effects estimate comparisons, shown in table 36, confirmed that VR meetings were rated significantly lower than both MS Teams and F2F ($p < 0,05$), while the difference between MS Teams and F2F remained non-significant ($p > 0,05$).

Person	(Intercept)
1	−0.947
2	0.258
3	1.062
4	−0.545
5	−0.143
6	−0.143
7	0.459

Table 37: random effects estimates table for meeting type.

The random effects table, shown in table 37, provides insights into the variance attributed to individual differences among participants. In this model, the random intercept accounts for participant-specific rating tendencies. These results indicate small but notable variability

among participants, confirming that individual rating tendencies exist but do not overshadow the effect of meeting type. This variation is accounted for in the model, ensuring that meeting type differences are not due to individual rating tendencies but reflect actual differences between VR, MS Teams, and F2F.

10.4.2 EMP

Effect	df	F	p
MT	2, 12.00	11.355	0.002

Note. Model terms tested with Satterthwaite testMethod.

Table 38: results from the EMP model.

Term	Estimate	SE	df	t	p
Intercept	7.810	0.397	6.000	19.649	< .001
MT (1)	−1.381	0.302	12.000	−4.566	< .001
MT (2)	0.333	0.302	12.000	1.102	0.292

Table 39: fixed effects estimates for the EMP.

Person	(Intercept)
1	−0.338
2	1.556
3	0.135
4	−0.102
5	−0.812
6	−0.338
7	−0.102

Table 40: random effect estimates for the EMP.

10.4.3 DSP

Effect	df	F	p
MT	2, 12.00	21.000	< .001

Note. Model terms tested with Satterthwaite testMethod.

Table 41: results from the DSP model.

Fixed Effects Estimates

Term	Estimate	SE	df	t	p
Intercept	7.238	0.416	6.000	17.397	< .001
MT (1)	−1.952	0.371	12.000	−5.264	< .001
MT (2)	−0.238	0.371	12.000	−0.642	0.533

Table 42: fixed effects estimates for the DSP.

Person: Random Effect Estimates

Person	(Intercept)
1	−0.947
2	0.258
3	1.062
4	−0.545
5	−0.143
6	−0.143
7	0.459

Table 43: random effect estimates for the DSP.

10.5 Appendix E

The observations per theme and per meeting type are presented here.

10.5.1 VR observations

Non-verbal cues

Most participants turned their head towards the speaker, two participants were all over the place.

Participants mostly do not maintain eye contact, they switch from looking at the PowerPoint presentation to the speaker

Participants at first did not use hand gestures but later into the session they did. Participants mostly look towards the PowerPoint presentation.

The active speaker turns his body towards the PowerPoint presentation and then back to the crowd often. The rest of the participants make a lot of movements with their arms but it has no clear goal. There are little to no head movements.

Additional notes: the active speaker moves his arms a lot, also rotates from the crowd to the PowerPoint. The crowd does not move a lot. They have an active listening posture but there is also chaos sometimes.

Engagement

The active speaker is the only speaker, there was no turn taking.

There was focus on the active speaker after the first 5 minutes. During the first 5 minutes some participants were busy with walking around trying to get into a chair.

There was no multitasking during the meeting, everyone was actively engaged into the presentation.

Collaboration

There was no turn-taking, there was one instance where someone was having issues so the observer had to help and his sound interrupted the presentation.

The active speaker addressed one person by name

The only collaborative tool used was the PowerPoint presentation.

Technical issues

There were no audio or video issues.

There were some participants who had some struggles with movement and getting into their chair.

There were no instances with motion sickness observed.

10.5.2 MS Teams observations

Non-verbal cues

Most of the time more than half of the participants are not looking at the screen. Most participants leaned their heads forward but not their eyes at the screen.

The active speaker used hand gestures once but they could not fully be seen by the crowd.

Participants were sometimes actively nodding.

Some participants were leaning forward showing attentions, while others were slouching indicating disengagement.

One observer noted “I see little signs of interest”.

Engagement

There was only one active speaker, but a few questions were asked.

Participants had their eyes often somewhere else than focused on the screen.

There were two instances where multitasking was observed by both observers.

Some participants were looking around, and showing inattentiveness.

Collaboration

There was only one active speaker.

The active speaker called people by their name when they addressed them.

The only collaborative tool used was the PowerPoint presentation.

Technical issues

There were no issues at all, everything worked seamlessly.

10.5.3 F2F observations

Non-verbal cues

Most people were switching between looking at the PowerPoint presentation and the active speaker. When a slide was longer visible people tend to look more toward the active speaker. Also when less text was on the slide.

People sometimes maintained eye contact.

Gestures were actively used during speaking.

There was a lot of nodding, especially when the active speaker was having eye contact with someone, then that person often nodded. More engagement was observed.

Most participants were leaning forward showing attention.

General notes: active listening pose, people look mostly towards the PowerPoint but nodding was used a lot.

Engagement

There were two active speakers, most participants contributed verbally to the conversation.

Most participants were focused on the active speaker, sometimes people looked towards their hands.

There was no multitasking observed.

There were a few signs of disengagement but this was very little, fidgeting.

General notes: There was more laughing and jokes than during VR/MS Teams.

Collaboration

There was a smooth and orderly turn-taking order. They asked each other for additional information and seemed to help each other, felt smooth

No-one called each other by name.

Only the PowerPoint presentation was used as active collaborative tool

There was one instance where others assisted in problem-solving, someone couldn't find the English word for self-checkout so someone helped.

Technical issues

No issues here

Instructions: Please rate your experience on the following questions. One answer per question.

Face-to-face Questionnaire	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree
Co-presence					
I noticed my team members.	O	O	O	O	O
My team members noticed me.	O	O	O	O	O
My team members presence was obvious to me.	O	O	O	O	O
Attentional allocation					
I was easily distracted from my team members when other things were going on.	O	O	O	O	O
I did not receive my team members' full attention.	O	O	O	O	O
I remained focused on my team members throughout the meeting.	O	O	O	O	O
Perceived message understanding					
It was easy to understand my team members.	O	O	O	O	O
My team members found it easy to understand me.	O	O	O	O	O
My team members' thoughts were clear to me.	O	O	O	O	O
Perceived affective understanding					
I could tell how my team members felt.	O	O	O	O	O
My team members emotions were not clear to me.	O	O	O	O	O
My team members could tell how I felt.	O	O	O	O	O
Perceived emotional interdependence					
I was sometimes influenced by my team members' moods.	O	O	O	O	O
My team members' feelings influenced the mood of our interaction.	O	O	O	O	O
My team members' attitudes influenced how I felt.	O	O	O	O	O
Perceived behavioral interdependence					
My team members' behaviors were influenced by my actions.	O	O	O	O	O
I reciprocated my team members' actions.	O	O	O	O	O
My team members behavior was closely tied to my behavior.	O	O	O	O	O

What grade 1-10 (1 is the lowest score, 10 is the highest score) would you give **Face-to-face** for the following:

1. The degree of perceived social presence. |____|
2. The effectiveness for this meeting purpose. |____|

Instructions: Please rate your experience on the following questions. One answer per question.

Microsoft Teams Questionnaire	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree
Co-presence					
I noticed my team members.	O	O	O	O	O
My team members noticed me.	O	O	O	O	O
My team members presence was obvious to me.	O	O	O	O	O
Attentional allocation					
I was easily distracted from my team members when other things were going on.	O	O	O	O	O
I did not receive my team members' full attention.	O	O	O	O	O
I remained focused on my team members throughout the meeting.	O	O	O	O	O
Perceived message understanding					
It was easy to understand my team members.	O	O	O	O	O
My team members found it easy to understand me.	O	O	O	O	O
My team members' thoughts were clear to me.	O	O	O	O	O
Perceived affective understanding					
I could tell how my team members felt.	O	O	O	O	O
My team members emotions were not clear to me.	O	O	O	O	O
My team members could tell how I felt.	O	O	O	O	O
Perceived emotional interdependence					
I was sometimes influenced by my team members' moods.	O	O	O	O	O
My team members' feelings influenced the mood of our interaction.	O	O	O	O	O
My team members' attitudes influenced how I felt.	O	O	O	O	O
Perceived behavioral interdependence					
My team members' behaviors were influenced by my actions.	O	O	O	O	O
I reciprocated my team members' actions.	O	O	O	O	O
My team members behavior was closely tied to my behavior.	O	O	O	O	O

What grade 1-10 (1 is the lowest score, 10 is the highest score) would you give **Microsoft Teams** for the following:

1. The degree of perceived social presence. |____|
2. The effectiveness for this meeting purpose. |____|

Instructions: Please rate your experience on the following questions. One answer per question.

Virtual Reality Questionnaire	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree
Co-presence					
I noticed my team members.	O	O	O	O	O
My team members noticed me.	O	O	O	O	O
My team members presence was obvious to me.	O	O	O	O	O
Attentional allocation					
I was easily distracted from my team members when other things were going on.	O	O	O	O	O
I did not receive my team members' full attention.	O	O	O	O	O
I remained focused on my team members throughout the meeting.	O	O	O	O	O
Perceived message understanding					
It was easy to understand my team members.	O	O	O	O	O
My team members found it easy to understand me.	O	O	O	O	O
My team members' thoughts were clear to me.	O	O	O	O	O
Perceived affective understanding					
I could tell how my team members felt.	O	O	O	O	O
My team members emotions were not clear to me.	O	O	O	O	O
My team members could tell how I felt.	O	O	O	O	O
Perceived emotional interdependence					
I was sometimes influenced by my team members' moods.	O	O	O	O	O
My team members' feelings influenced the mood of our interaction.	O	O	O	O	O
My team members' attitudes influenced how I felt.	O	O	O	O	O
Perceived behavioral interdependence					
My team members' behaviors were influenced by my actions.	O	O	O	O	O
I reciprocated my team members' actions.	O	O	O	O	O
My team members behavior was closely tied to my behavior.	O	O	O	O	O

What grade 1-10 (1 is the lowest score, 10 is the highest score) would you give **VR** for the following:

1. The degree of perceived social presence. |____|
2. The effectiveness for this meeting purpose. |____|

10.7 Appendix G

Observation Template for Experiment

Experiment Details

Observer Name: _____

Meeting Modality: ☐ VR ☐ Microsoft Teams ☐ Face-to-Face

Discussion Topic: _____

Duration of Discussion: 15 minutes

Non-Verbal Cues & Interactions

Observing how participants use body language, gestures, and avatar movements to engage with the group.

Behavior	Description & Notes
Head Orientation	Are participants turning their heads toward the speaker? ☐ Yes ☐ No
Eye Contact (F2F only)	Do participants maintain eye contact (or simulated avatar gaze in VR)? ☐ Yes ☐ No
Gestures While Speaking	Are participants using hand gestures to emphasize points? ☐ Yes ☐ No
Avatar/Body Movement (VR)	Are participants moving around, nodding, or using body posture in VR? ☐ Yes ☐ No
Seating/Posture	Are participants leaning forward (showing attention) or slouching (indicating disengagement)? ☐ Yes ☐ No

Additional Notes:

Engagement Levels

Tracking signs of attention, participation, and focus during the discussion.

Behavior	Description & Notes
Speaking Turns	Do all participants contribute verbally? ☐ All ☐ Most ☐ Few ☐ None
Attention to Speaker	Are participants focused on the active ? ☐ Yes ☐ No
Multitasking (Teams Only)	Are participants checking emails, using phones, or doing other tasks? ☐ Yes ☐ No
Signs of Disengagement	Are participants looking around, fidgeting, or showing inattentiveness? ☐ Yes ☐ No

Additional Notes:

Collaboration Quality

Observing how participants interact, build on each other's ideas, and work together during the discussion.

Behavior	Description & Notes
Turn-Taking	Do participants speak in an orderly manner, or do they interrupt frequently? ☐ Orderly ☐ Frequent interruptions
Direct Engagement	Do participants directly address one another by name? ☐ Yes ☐ No
Use of Collaborative Tools	Are tools such as whiteboards, sticky notes, or VR objects being used? ☐ Yes ☐ No
Helping Others	Are participants assisting others in problem-solving or building on ideas? ☐ Yes ☐ No

Additional Notes:

Technical Challenges (VR & Teams)

Tracking usability issues, technical disruptions, and participant comfort.

Issue	Description & Notes
Audio Issues	Any microphone/sound problems? ☐ Yes ☐ No
Lag & Video Quality (Teams & VR)	Any delays, stuttering, or crashes? ☐ Yes ☐ No
Usability Problems (VR Only	Are participants struggling with controls or movement? ☐ Yes ☐ No
Motion Sickness (VR Only)	Are any participants showing discomfort or dizziness? ☐ Yes ☐ No

Additional Notes:

Additional Observations & Researcher Notes