The Impact of Avatar Fidelity on Virtual Reality Brainstorming and Room Design Evaluation

Thesis Project for the Master of Science in Human Computer Interaction

Utrecht University

Author: John Scott Siri Jr.

Student ID: 4871030

Supervisor: Dr. Wolfgang Hürst

Second Examiner: Dr. Zerrin Yumak

External Supervisor (TNO): Jelle Wolbers MSc

Second Supervisor (TNO): Dr. Jonathan Barnhoorn

Date: June 17th, 2024

Location: Utrecht, Netherlands

Abstract

Multi-User Virtual Reality has potential applications for prototyping environments, but questions still exist about how user representations influence brainstorming and evaluation activities. Improvements in avatar fidelity have led towards avatar realism being more commonplace, but priming from avatars may influence users to behave differently. This study examines the differences in brainstorming activities between High and Low states of Avatar Fidelity while in Multi-User VR workspaces. Using kitchens as a workspace everyone has worked in and can comment on, groups of participants went through sets of kitchen prototypes and came up with Pros, Cons, and Recommendations for each. Participants were given surveys where they reported information about team dynamics, perceptions of the environments evaluated, behaviors while brainstorming, and environmental preferences.

After performing the study, significant quantitative differences were not found regarding current state of Avatar Fidelity using paired *t*-tests. However, exploratory analyses using independent *t*-tests for each Room Set found significant differences in team dynamics and behaviors while brainstorming between those starting in High Fidelity and Low Fidelity. Additionally, qualitative analyses of obversions and open ended questions indicated differences in focus relating to stylization, appliances, and environmental interactivity. These variations might be contextually relevant for different use cases of brainstorming in Multi-User Virtual Reality.

Acknowledgements

This Master's Thesis has been in the works for a long time and is the culmination of a great deal of thought and effort. There are many people I would like to thank for helping me get through this endeavor:

I would like to thank my university supervisor, Wolfgang Hürst, for being my tether to the university during my time spent working on the thesis. I would also like to thank my second examiner, Zerrin Yumak, for taking the time out of her schedule to act as a member my examination team.

I would like to thank my TNO supervisors, Jelle Wolbers and Jonathan Barnhoorn, for all their guidance and insight. They kept me running on the day-to-day basis and always made time for me. The both of them really helped me through each step of my thesis journey. This thesis would not have been nearly as robust if not for their input and encouragement.

I would also like to thank the folks at the TNO Human-Machine Teaming department. I was surrounded by many talented and experienced people who made me feel very welcome while I worked alongside them. They gave me the physical and mental resources to make this project possible.

Lastly, but not least, I would like to thank all my friends and family for encouraging me and keeping me from working myself to the bone. Their support means the world to me and I am lucky to have them by my side (even if an ocean apart).

Table of Contents

1 Introduction

In the past decade, there have been numerous developments in the multi-user virtual reality space. One area in multi-user virtual reality that has received a lot of focus is regarding serious gaming, where users play games for reasons other than entertainment. While often used for education and training, virtual reality has potential in the realm of concept development and prototyping. Virtual reality can allow teams to demonstrate designs in different states of depth and directly reference relevant elements in a fairly immersive 3-dimensional space. Spatial prototypes can potentially allow the concept iteration and refinement processes to happen with more substantial feedback, having achieved a better level of understanding through spatial understanding, before major resources get invested into a developing idea. Users engaging with virtual reality as a design tool can possibly identify problems and share a cohesive idea on what a design concept may look like with insights more that are more specific and clear than attained with 2-dimensional schematics. This is particularly relevant in the defense and industrial sectors, in which these processes may occur over extended timeframes and the ability to experiment and iterate early on in the process is beneficial. At TNO Defense, Safety, and Security in particular, the Human-Machine Teaming department is interested in developing room evaluation methods for virtual space prototypes in virtual reality. Constructing spaces can be expensive and problems not caught can impact effectiveness of workspaces, so being able to explore special prototypes in virtual reality is a way to test environments before resources are poured into building them.

However, a significant factor in evaluating designs and concepts is the quality and depth of constructive feedback given. Design thinking strategies, like brainstorming and rapid prototyping, work to increase openness to new ideas and help iterate through many possibilities, but they are largely performed externally and before a substantial prototype is developed. In a virtual reality environment, there is opportunity for flexible 3-dimensional drafting in earlier concept development stages with lower stakes while maintaining the benefits to openness (Minas et al., 2016). This is an indication for exploring innovative ideas and manifesting potentially rich insights that are oftentimes hard to demonstrate or explain until there is a more tangible existing design present. Furthermore, a 3-dimensional design space is capable of helping establish a united idea between all existing members within a team and can prevent misunderstandings.

Partially due to the flexibility a simulation provides for interaction, there is the potential for distractions. Some of these may have to do with the visualization of the environment, but avatars themselves present an influence. Displaying different appearances may change the perception of other users, their ideas, and their willingness to participate or interact. As of late, there has been developments regarding different levels of realism applied to avatars and their environments. Higher fidelity avatars have been shown to have increase presence and embodiment and are often capable of displaying social cues in higher fidelity (Aseeri & Interrante, 2021). However, early stage design prototypes may vary in visual fidelity, creating the opportunity for a lack of synchronization between the environmental tone and avatar visualizations. Further, the avatar designs may possibly impact the view may change their engagement with teammates and level of skepticism they view their input with.

In this study, we examine the impact different levels of avatar visual fidelity have on the feedback provided in the multi-user prototyping environment. Based on current research, there is a chance that different avatar visualizations may prime certain attitudes towards interaction and responses. There is also a chance for disengagement when dealing with a lack of cohesion between the avatar appearances and the environmental fidelity. This would be useful to know when exploring concept development and evaluation, as it may indicate avatar considerations when trying to stimulate feedback.

2 Literature Review

In order to look into how multi-user VR can be used for brainstorming in room design, there is a need to understand the existing research. Understanding the way avatars are perceived by their user and others in the environments is a necessity for developing multi-user VR systems of any kind. Awareness of collaboration factors is important for knowing behaviors and effects to look for in situations where people need to work together in VR. Additionally, knowledge on the current examples where VR is being used as a design tool can tell much about how related systems have been and could be used when doing brainstorming activities.

2.1 Avatar Perception

2.1.1 Avatars & Visual Fidelity

Appearances in virtual reality are flexible. Users are capable of embodying avatars that can be in different states of visual abstraction and contain different physical characteristics. While avatar presentation is often thought of as an outlet for a user to represent themselves, the avatars used can alter the user's cognition and change the ways they think about both themselves and their experiences. Individually, the differences in avatar representation may impact their frame of reference or their own self-perception. The visual fidelity of the avatar and its presence has been shown to impact user's judgment of egocentric distance, where users with simplified avatars and no visible avatar were similarly and significantly worse at these tasks (Ries et al., 2009). These consequences are not negligible when considering the ways that users evaluate their environments, because it means that their own avatars impact their understanding of virtual spaces. Realistic avatars have also been found to be more human-like, have demonstrated higher levels of virtual body ownership, and that other's avatars impacts a user's own self-perception (Latoschik et al., 2017).

However, the impacts of avatar perception are not black and white. The consequences of avatar fidelity have been previously shown in social situations, though the specific implications are still unclear. In social circumstances, data collected from surveys suggests that limited avatars consisting of heads and hands do not have significantly different copresence and behavioral interdependence scores from avatars consisting of complete bodies (Heidicker et al., 2017). It has also been found that reduced avatar fidelity does not significantly impact bystanders ability to interpret interaction behavior in virtual reality, even when avatars are very simple (Mathis et al., 2021). This would imply that outsiders are capable of identifying mechanical interactions between users regardless of the realism of the avatar. Further yet, there are indications that visual realism alone does not lead to significant differences in empathy towards virtual characters (Higgins et al., 2023). These would indicate that a lower fidelity avatar is fairly sufficient for many social circumstances and indicate that there is a lower social need for higher fidelity avatars. It should be noted that low fidelity avatars are sometimes portrayed as similar to cartoons and other times floating heads with hands or stick men and virtual reality technologies have been improving quickly. These differences in impact mean that while avatar representations have consequences, these consequences are not unilaterally linear on a broad scale and differ on application. It also means that avatar representations may have unique impacts dependent on the situation they are used in.

Avatars that are of higher fidelity and considered more realistic tend to be rated as more engaging than lower fidelity avatars or those considered more cartoony or abstract (Roth et al., 2016) (Latoschik et al., 2017). However, attempts at realism may cross over into the uncanny valley, creating a relatively undesirable sense of discomfort at a partially realistic but 'off' avatar (Higgins et al., 2023). This is important to consider given that while many existing services that exist use primarily cartoonish avatars, while there have been developments in the fidelity of avatars and a general trend toward increased realism in virtual reality. This does not always bode well; attempts to increase realism to increase visual information from an avatar fall flat when the avatar is viewed as eerie (Shin et al., 2019). There have since been attempts to ‘cross the uncanny valley’ wherein photorealistic avatars are projected using advanced visual techniques and are able to display emotional states in a more realistic way (Higgins et al., 2021). This has resulted in a general upward trend in the metrics of social presence and immersion (Roth et al., 2016). With this in mind, there is little evidence at the current moment that avatars that are more realistic are better when considering task load or performance, especially in group settings. Confounding factors exist as well, where users are biased towards finding the avatar they used first in a set of trials as less creepy than others used, regardless of realism (Ma & Pan, 2022). This bias is not entirely understood at the moment and it may influence how avatars frame experiences. This indicates that there are likely circumstances where cartoonish characters may not be inherently inferior to the higher fidelity avatars.

2.1.2 Avatars, Cognition, & Context

Avatars help define people’s interactions in virtual spaces. Priming, where individuals encounter stimuli which prompts existing knowledge structures based on situational context, happens automatically and can influence perceptions and behavior (Bargh et al., 1996). The Proteus Effect, where online representations of users have been found to impact behavior and self-perception, is an example of the impact of priming and can influence social interactions that occur online (Peña et al., 2009). This behavior has been shown to be capable of impacting they way players treat each other in teams (Sengün et al., 2022). Users in different virtual environments have demonstrated preferences for different types of avatars depending on the context they were in, with low-poly avatars being preferred in work contexts and realistic avatars being preferred in casual social contexts (Praetorius et al., 2021). This can be similar in classroom settings too, where there are indications that users with avatars with similar to the user's external appearance demonstrate higher self-presence and lower enjoyment when compared to users with uniform avatars (Han et al., 2022).

Avatars and environments have more overlap with each other than might be expected. When using avatars that look more like themselves, users reported that they found other user's avatars and the environment quality to be more realistic and displayed more synchronous nonverbal behaviors with other users (Han et al., 2023). This same study found that users in outdoor virtual environments and more spacious virtual environments indicated scores for higher self-presence, spatial presence, and perceived realism (Han et al., 2023). These influences are important, as they highlight a relationship where the environment, avatars, and cognition all prime the user in some way. In this study, it is relevant to determine the effects priming from different avatars have on the evaluations about the environment and the ideas that come from it.

2.1.3 Avatar Effects on Other Users

Not all avatars are viewed the same way. Users may display preferential treatment to avatars that meet certain criteria. There has been evidence that avatars with moderate similarity to a user's own avatar are more easily persuaded and that said user finds both avatars with identical and similar facial shapes easier to interact with than avatars with dissimilar facial shape (Shih et al., 2023). This difference in capacity to persuade indicates that the visualization of an avatar is capable of changing communication results between each other. When concerning teams, an avatar's capacity to persuade is a factor to consider because differences in team communication can intermingle with team's capacity to perform effectively. However, these studies still do not examine the impact fidelity of said avatars has on willingness to engage with one another and work together. Simpler avatars may potentially be more similar to one another due to the few characteristics present and therefore fewer ways to be different. This knowledge is important for collaborative virtual reality because it implies that the avatars are capable of priming reactions from users and lead to different interaction results.

There have also been previous findings that copresence is stronger for more realistic avatars (Casanueva & Blake, 2001). Co-presence was also found to impact navigation strategies in interior virtual environments (Yassin et al., 2021), which is relevant to the ways teams might explore room prototypes. Furthermore, there is some proof that gesture is an important factor in computer-mediated brainstorming environments as a creativity support tool (Liao & Wang, 2019).

When working directly with other people, the way other users are represented alone can impact the way tasks are performed. In Kocur et al. (2020), two players completed the Tower of London cognitive task, first alone then together. The control group for this had two generic avatars while the test group had one of the avatars resembling Albert Einstein. Interestingly, the presence of the virtual Einstein caused a significant difference in perceived workload and cognitive performance after performing the task for the other player when compared to the control group and the virtual Einstein himself. This indicates that the avatars of other players near you can impact your own cognitive performance and perceived workload depending on the association with the avatar portrayed (Kocur et al., 2020). Additionally, this demonstrates tangible overlaps between avatar visualization and task execution because this explicitly demonstrates that the visualization of others can impact a user enough to significantly affect their mental state when performing an identical task.

2.2 Collaboration Factors in Virtual reality

2.2.1 Communication Rules

Collaboration with others is built on fundamental communication rules. The need for communication persists wherever there is a need for coordination and collaboration. To this end, it is required to communicate effectively when evaluating designs, as it is an inherently iterative experience that requires feedback. The communication required for effective design evaluation and brainstorming are not exclusive to the real world, but also within virtual reality. However, there are challenges for team communication in virtual reality, especially when there is a good deal of difference between the avatars and their features.

It has long been noticed that many forms of digital communication carry difficulties in conveying complex information compared to in-person communication methods (Straus & McGrath, 1994). Cue-poor environments require users rely on fewer cues to orient themselves and conversation becomes less spontaneous. However, cue-rich environments have the potential to produce more naturalistic interactions. Avatar design is capable of influencing this. For example, sufficient embodiment in virtual reality was found to result in similar communication patterns to face-to-face interactions and higher levels of social presence when compared to users with no-embodiment, which were lonely and suffered from worse communication (Smith & Neff, 2018).

After a conversation starts, there is a need to regulate the flow of conversation amongst its participants. While this is the case is dyads, triads and bigger groups have complexities in understanding conversational direction. Behaviors like turn-taking and addressing are thereby used as collective discussion management tactics. Turn-taking behaviors are the local management tactic for determining who speaks when throughout the course of an exchange (Sacks et al., 1974). These behaviors occur constantly throughout conversations and can be found in virtual environments as well (Jenks, 2009). Turn-taking contains series of verbal and nonverbal cues that work together and simultaneously help indicate turn order.

Group communication is often more complicated than between a pair of participants. The bigger a team gets, the more people need to coordinate together in order to stay on the same page. When in triads, communication between two members is influenced by the third being present (Butts, C. T., 2008). There are more possible interactions with different sets of individuals when more individuals are present, which results in increased interaction complexity. Addressing is a common strategy for directing communication towards a given individual, subgroup, or topic. Addressing takes place through a combination of speech, gaze, and gesture and may demand auditory and visual attention respectively (Jovanovic et al., 2006). It has been found that group members look towards each other more frequently when listening rather than when speaking (Maran et al., 2021). Strategies regarding addressing allow participants to effectively discuss relevant topics with relevant participants. This tactic is particularly common when used with larger groups because not everyone has the same information available. Within virtual spaces, these tactics are also used to manage discourse, though the presence or absence of nonverbal cues can influence the discussion.

Nonverbal cues, like gaze and gesture, are core parts of natural communication. Gaze is used as a regulator within conversational flow both inside virtual reality and externally. Mimicry of posture and position tend to enhance factors like persuasiveness (Van Swol, 2003). Gaze impacts the dynamics of interruptions, which can ensure that communication is clear between involved parties. Gaze signals visual attention, which can be used to ensure clear communication. This is vital to note for the purposes of this study due to the clear and direct link between environmental evaluations and user's capacity to clearly signal topics in the virtual environment while conversing.

2.2.2 Avatar-based Expression

Avatars are an outlet for users to communicate with each other through nonverbal means. This can occur passively or actively, but it is an integral part of communication with each other. Considering that nonverbal cues have significant visual components, this makes adequate representation an important feature when considering building virtual reality systems. This has led to a longstanding push to increase the capacity of avatars to convey emotional expression, especially with their faces (Basori & Ali, 2013). Sufficiently immersive virtual humans are capable of training social conversational protocols better than guides comprised of illustration and explanation (Babu et al., 2007). A 2020 study implied that facial expressions were more important than body movements in conversational outcomes (Oh Kruzic et al., 2020). These indications imply a potential need for a higher avatar fidelity in order for effective social interaction.

In Tanenbaum et al. (2020), players performed a series of five challenges in virtual reality, each with a different partner. Three of the tasks were primarily cognitive and two were considered more physical. While doing this, the players were able to manually shift emotional visualizations on their avatars to represent their emotional state. However, the manual sorting of emotions was infrequently used due to its manual nature and primarily relied on auditory cues instead. The ease of expression is thus something further considerable, as there are assortments of methods and control schemes for actually controlling the emotional expression (Tanenbaum et al., 2020).

This push for expressiveness is perhaps not wholly the same as an avatar's realism as a study involving both realistic and cartoon-like characters noted that participants found cartoon-like character's facial expressions easier to control (Ma & Pan, 2022). This somewhat deviates from the push to always develop a more realistic avatar as new technology emerges that is capable of plausibly rendering it. Even if valuable, users tend to neglect inconvenient displays for self-expression in social virtual reality (Khojasteh & Won, 2021). This does not mean that emotional expressions are not valued by users though, as the absence of facial animations for conveying emotions was a pain point in simulated co-located multiplayer experiences (Sykownik et al., 2023).

Additionally, eye representation, and gaze derived from it, has been noted as a key perceptual feature. Developments in eye tracking have led to improvements in virtual reality platform's capacity to display eye rendering and positioning, which can impact impression formation (Weibel et al., 2010). Significantly, there are indications that quality and realism of gaze is related to the amount of eye contact given in virtual settings (Krum et al., 2016). In remote mixed reality collaboration, gaze cues and gesture cues lead to significantly higher copresence than a control without and a combination of gaze and hand gesture cues were found to both significantly increase co-presence than gaze alone when conveying spatial actions and demonstrate a significant difference in task completion time when comparing the combined group to the control (Bai et al., 2020). Furthermore, gaze has been shown to play a significant role in moderating turn-taking behaviors in casual conversational interactions (Jokinen et al., 2013). This indicates that there is likely a role that gaze would play explicitly in interactions between users and helps function as a tool in conversational awareness between conversing users, which can impact the interactions between team members.

Despite the benefits that avatar expression is capable of providing, cases exist where displaying avatar features is not preferred or is negligible regarding performance. When implementing blinking, methods were found that increased distraction and noticeability but not game performance (Zenner et al., 2023). Additionally, eye-tracking was found to be insignificant to the quality of dyadic virtual reality conversations even though it was considered more realistic than simulated gaze (Andrei et al., 2023). Even though eyes have long been recognized as important social features, these eye-related features did not improve information across different contexts. Other attempts to visualize gaze have also been developed. In one such attempt, a pair of users with visible representations of their vision (depicted as oval outlines) were overlayed around where a user was looking in a data comparison task, but no significant effect on performance was found and users in control conditions were found to use hands as indicators of attention when the visualizations were not available (Bovo et al., 2022).

Bodily positioning between interacting partners has significant value in establishing communication between two users. The combination of bodily and facial features on avatars may be more effective when used together than either feature alone (Sidenmark & Gellersen, 2019). These related and often coinciding cue interfaces play a significant role in nonverbal communication, which may affect the ways that teams may make decisions. During the performance of the vision cone and data exploration study, users who did not have the vision cone enabled compensated via using facial proximity and hands to indicate where they were looking at or referencing (Bovo et al., 2022). This is interesting itself, as this proposes the notion that users are capable of mitigating the absences of some facial cues with bodily cues and may use their body to perform additional clarifications in situations where level of comprehension is difficult to understand. This creates an interesting scenario, where alignment of different cues leads to more comfortable and effective communication, but users are able to compensate for cue absence with other indicators in certain situations. These deviations from communication expectations make it difficult to gauge the exact level of potency features like eye tracking provide to virtual reality users.

2.2.3 Engagement in Creative Activities

Teams of designers are required to engage in brainstorming and critiquing activities as part of their role. This means that openness to new ideas and critique are essential parts of the collaborative design experience. However, engagement and participation in group brainstorming and revision activities is not always effective or equal across the board. Restrictions on communication like production blocking, where teammates need to take turns to express their own ideas, inhibit brainstorming productivity (Diehl & Stroebe, 1987). Encouraging proactivity and engagement behaviors while doing group activities allows for more opportunities for construction and revision. It was found that even simple electronic brainstorming activities can lead to reductions in production blocking, especially in larger groups (Gallupe et al., 1992). When performing design thinking activities as dyads online, sets with two highly motivated teammates were found to be the most efficient communicators and achieve the best creative results, though had less overall interaction with each other than groups with one high motivation teammate and one low motivation teammate (Zhang et al., 2022). The willingness to engage in the creative task is an important influence to the output provided.

It has been found that socially anxious individuals are less productive in group brainstorming activities, possibly because of the nature of group interaction (Camacho & Paulus, 1995). However, virtual tools have been shown the capacity to somewhat mitigate these effects. On a similar note, shy people felt less communication apprehension in virtual reality compared to face-to-face communication, but virtual reality was found to have less persuasive power than face-to-face interaction (Hammick & Lee, 2014). This indicates the possibility for virtual reality to encourage higher levels of participation than in-person brainstorming and evaluation activities, though perhaps makes users more likely to disagree with one another.

Auditory and visual attention have different effects on cognition in in virtual reality. Nonverbal cues are primarily visual and nature and the core of conversations themselves focus on the auditory aspects of communication. In virtual reality, tasks that are in these different modalities lead to different cognitive effects on the involved participants (Voinescu et al., 2020). One such consequence of this is that visual stimuli tend to increase impulsivity and auditory stimuli tend to increase inattention (Voinescu et al., 2020). It was also found that oral communication facilitates better creative output than written modality (Forens et al., 2015). These differences in cognition and output based on modality perhaps mean potentially substantive impact when considering tasks that swap between focuses in modality. Design evaluations, for example, have a visual component to actually understand a situation spatially and make judgements, but require auditory components to discuss with other participants.

Criticism is a major factor of the iterative revision process. In the evaluation sense, being able to point out problems with the current way things are designed is how change gets made. However, criticism is a touchy subject and can have negative repercussions when the tone gives off the wrong message. Individuals who receive destructive tend to avoid or clash with the source of criticism instead of collaborate or compromise more than those who received constructive criticism and is a substantial cause of conflict (Baron, 1988). While brainstorming, criticism is often considered to be against productive brainstorming and is discouraged (*What Is Brainstorming?*, 2016). Despite criticism being considered more divisive while brainstorming in competitive social contexts, there is some indication that suggests that encouraging criticism while brainstorming in cooperative social constructs increases overall productivity and is generally taken positively (Curhan et al., 2021). Nevertheless, criticism is a potentially risky type of interaction that can impact the group dynamic and discourage the free flow of ideas if done in the wrong way or the wrong context.

2.3 Virtual Reality as a Design Tool

2.3.1 Use Cases & Limitations

Design thinking has become a core construct to the development of ideas through different brainstorming, evaluation, and iteration mechanisms. Design software like Figma has become an essential pillar regarding the development of software goods and services. Tools, tactics, and practices are commonly experiences when undergoing technological development. Designs for physical and even environmental products go through a drafting process using software like AutoCAD. However, while these 3-dimensional plans are often highly detailed, they are still viewed through spatially limited by a 2-dimensional interface and often are focused on precision rather than iteration and immersion. When brainstorming and evaluating designs in a 3-dimensional space, virtual reality is a possible means for compensating for these gaps in the design process and allow for immersive design practices to take place. To this end, there have been many use cases for design techniques to be implemented in virtual reality by way of immersive service prototyping that can often provide more insight than conventional methods (Abdel Razek et al., 2018). If these methods are as effective as they are promising, then virtual reality has the opportunity to become an incredible tool within a myriad of fields where design is an essential step.

The belief that virtual reality is a tool capable of enhancing the creative experience is not unfounded. Virtual reality has been utilized as a design tool for interaction designers, who noted that the virtual environment provided unique design explorations resulting from the space (Jetter et al., 2020). Virtual environments were demonstrated to have a clear impact on creativity, though recreating the original brainstorming environment in virtual reality is not as positively impactful and implies that the context of the virtual environment is important (Guegan et al., 2017). Virtual reality has been also applied to remote scientific collaboration, which had their own series of design guidelines for supporting their needs (Olaosebikan et al., 2022). The scientific process is fairly structured, but also requires brainstorming aspects and the ability to discuss different ideas, showcasing virtual reality's utility as a powerful tool for coordinating academic insight remotely. These examples demonstrate the potential that virtual reality design applications are capable of and that they are an area being explored. However, these design implications carry different weight for designing experiments or performing interaction design instead of room design.

Service designers have noted that virtual reality is capable of demonstrating the ability to bridge the gap between prototype environments and the actual service environment, especially in service environments that take place in larger spaces or with human touchpoints present (Boletsis et al., 2017). Both large and with many human touchpoints throughout, virtual reality tools have already started being applied for room design in naval spaces in order to better bridge the gap from abstract idea to actual environment and provide feedback (Cassar et al., 2019). Additionally, building design has recognized the potential of virtual reality as a comparatively low cost, insightful step between pictures and construction (Paes & Irizarry, 2016). This is significantly related to the question in this study, as virtual reality is an emerging technology capable of helping professionals in both service and room design, though both are still refining their design recommendations for their own unique contexts in virtual reality.

However, as powerful as the tools can be can be, experts in service design have reported that they feel overwhelmed after using virtual reality prototyping for too long (Abdel Razek et al., 2018). Among other concerns, the scientists who tested collaborative virtual reality felt that avatars unable to sufficiently display social cues made it difficult to engage in turn taking behaviors and limited the effectiveness of communication in virtual reality (Olaosebikan et al., 2022). These are indicators that there are design decisions within the virtual reality platforms that are still limiting their performance in prototyping. The tools present in virtual reality appear to be potent, but still need to be further refined if they are to be used consistently and regularly.

Furthermore, there is not a specific linear set of tools that can be implemented in virtual reality for all use cases. The desired tools for collaboration in virtual reality need to be tailored to the tasks themselves in order to provide effective tools (Freeman et al., 2022). This lack of linearity prevents one-size-fits-all solutions from taking place and requires understanding of the domain that the virtual reality environment is being set up for.

2.3.2 Place in Frameworks

The demonstrated need for consistency between phases of development is important enough to be adapted into existing engineering structures. Design processes are often outlined by various structures to move the process along and keep the vision consistent between different phases of the project. One framework that is used to both enable creativity and moderate the differences that have emerged between ideas that have developed across the design process is TNO Defence, Security, and Safety's Concept Maturity Levels to communicate progress and give clarity across the development process (Van Der Wiel et al., 2010). In this framework, a common need arises to ensure that the different members of the team share the same vision when designing projects from robots to control rooms. This can be a particularly difficult case on the earlier levels of the design process, as it is still abstract and largely up to interpretation.

Historically, sketching and planning on paper has been an important facet in sharing these visions and is more effective than verbal discourse alone in providing shared understanding (Köping Olsson & Florin, 2011). However, spatial understanding is rather difficult when looking at sketches in 2-dimensional ways and there is still room for misunderstanding. Virtual reality provides a strong outlet with this space, enabling teams to evaluate environmental designs in a spatially comprehensible and uniform way while still early on in development. These virtual prototype designs can create consistency and clarity earlier in the process before where would taking an otherwise significant jump in labor invested.

3 Research Questions

3.1 Considerations

The differences in avatar perceptions mean that interpretation of tasks, environments, and both the other users' ideas and critiques are possibly able to be altered depending on the avatars used. This is important for both the user and their teammates, which are also influenced by the avatars displayed. When developing a system based around the notion of early stage iteration, it is important to ensure that the avatars present enhance the desired level of engagement on the right scale and scope. Existing research indicates the context matters to the ideal avatar, though differences required between domains in collaborative environments in development is still something not yet explored. Communication factors influence the way that users interact with each other and differences in cue availability is present with an avatar. An avatar's abilities to prime other users into different behavior demonstrates this importance in circumstances where team communication is paramount to the iterative design process. Abstract, lower fidelity avatars may inhibit normal bounds for communication, but may also be less bound by associations imparted by more realistic, higher fidelity avatars. Understanding which avatars lead to which effects in these early stages of environmental development would help with avatar selection for immersive, multi-user prototype environments.

There are many opportunities for virtual reality as a design tool for 3-dimensional spaces. While already being considered in service and room design, there still exists a need to refine the tools available to best fit the users. Avatars and their ability to communicate were indicated as a potential limitation for the effective use of design tools in virtual reality, but there is not clarity about the specific implications that avatars have in collaborative prototype environments as communication instruments.

3.2 Questions

The literature at hand identifies the opportunity to further explore the uses cases of multi-user virtual reality and the implications that avatar designs have on the overall process of design evaluation, brainstorming activities, and discussions. While there are many questions still unanswered, this study will focus on the effects different states of avatar fidelity have on room design brainstorming activities and evaluation feedback while in virtual reality. This prompts the following research question and sub-questions:

RQ1: When using multi-user virtual reality for evaluating room design, does the visual fidelity of the evaluators' avatars impact the feedback and ideas that they provide?

a. Does avatar fidelity impact team dynamics while doing evaluation tasks as a group?
b. Does avatar fidelity influence perceptions of the environment evaluated?
c. Does avatar fidelity impact behaviors while brainstorming?
d. Does avatar fidelity influence preferences after brainstorming?

3.3 Variables

In order to obtain measurements regarding the aforementioned research questions, variables needed to be assigned that are connected to the research questions. Using a set of surveys comprised of ad-hoc questions and excerpts from existing validated surveys, the following variables will be measured:

Table 1

List of Variables and Descriptions

RQ	Survey Source	Variable Name	Description
A	(Harms & Biocca, 2004)	Co-Presence	The degree to which the observer believes they are not alone and secluded, their level of peripheral or focal awareness of the others, and their sense of the degree to which the others are peripherally or focally aware of them.
		Perceived Message Understanding	The ability of the user to understand the message being received from the interactants as well as their perception of the interactant's level of message understanding.
		Attentional Allocation	The amount of attention the user allocates to and receives from other interactants.
		Perceived Affective Understanding	The user's ability to understand interactants' emotional and attitudinal states as well as their perception of the interactants' ability to understand the user's emotional and attitudinal states.
B	(Vorderer et al., 2004)	Spatial Situation Model	The user's mental model and awareness of the location they were just in.
		Self-Location	The amount a user views themselves as immersed in the virtual environment.
C	(Dennis & Valacich, 1993)	Satisfaction	The user's feeling that they are content with their contributions.
		Production Blocking	The user's sense that the limited ability to converse simultaneously suppressed idea generation.
		Evaluation Apprehension	The user's sense that they withheld ideas due to a fear of negative evaluation.
		Free Riding	The user's sense they relied on other to accomplish the task.
		Synergy & Stimulation	The user's sense that they were motivated to contribute because of others' participation.
		Sufficient Time	The user's sense that they had enough time to perform the task.
D	Ad-Hoc	Preference Ratings	The ratings each user gave regarding how much they liked the environment.
		Number of Pros	The number of pros that a user listed about the environment.
		Number of Cons	The number of cons that a user listed about the environment.
		Number of Recommendations	The number of recommendations that a user listed about the environment.

3.4 Hypotheses

With the literature in mind, I expect avatar fidelity to have a role in determining the specificity and scope of evaluator feedback. In particular, I expect that lower levels of avatar fidelity will lead to less specificity in feedback, but a larger scope. I think this because the users will have less cues to pay attention to, they will act less like a group when examining their environment, but also less consistency. I do think there is a risk that the low fidelity avatars will be somewhat distracting given that it will be harder for other evaluators to determine what they are doing. However, I also suspect that avatar fidelity might influence evaluator brainstorming due to the consequences of avatar-driven social cues and priming from the individuals.

Additionally, I expect that avatar fidelity has a role in priming users to willingness to engage in design discussions with each other. While there is still uncertainty regarding the exact situations that high fidelity avatars will be viewed as more useful than lower fidelity avatars, design spaces that feature high levels of communication imply a need for tools and indicators to passively facilitate communication. The environment present will be a combination of observation and communication, so the avatars with more observational cues will lead to increased openness and engagement in communication.

H1: Avatar fidelity will impact the feedback given by evaluators when evaluating room design together.

a. Team dynamics will be improved when the avatar fidelity is high.
    a. Co-Presence will increase when avatar fidelity is high.
    b. Perceived Message Understanding will increase when avatar fidelity is high.
    c. Attentional Allocation will increase when avatar fidelity is high.
    d. Perceived Affective Understanding increases when avatar fidelity is high.

b. Avatar fidelity will significantly influence perceptions of the environment.
    a. Spatial Situation Model will change significantly in different states of avatar fidelity.
    b. Self-Location will change significantly in different states of avatar fidelity.
c. There will be differences in brainstorming behaviors between the different states of avatar fidelity.
d. Avatar fidelity will not impact preferences while brainstorming.
    a. Avatar fidelity will not change the number of room advantages.
    b. Avatar fidelity will not change the number of room disadvantages.
    c. Avatar fidelity will not change the number of recommendations.
    d. Avatar fidelity will not significantly change the opinion on the room environments.

4 Methods

4.1 Study Design

The main factor is *avatar fidelity* (*avatar fidelity*: low vs. high). Participants would be grouped into groups of 3, each group will experience both avatar fidelity levels (within-groups design) over two sessions (*session* 1 and 2). In each round a set of two work environments (*work environment set*: A and B) will be evaluated. The order of avatar fidelity and work environment set will be counterbalanced based on participant number (see counterbalancing table in Appendix A). After each work environment and after each session, questionnaires will be administered. After round 2, a small interview will be held along with a short debriefing.

4.2 Participants

10 groups of 3 were planned, leading to a total of 30 slots. These slots were placed across 5 days, with 2 sessions per day (one in the morning and one in the afternoon). The following inclusion criteria needed to be met as a prerequisite for participation in this study: participants needed to be physically healthy (due to the need for motion and use of VR equipment) and needed to be proficient in English speaking, comprehension, reading, and writing (as I am not sufficiently proficient in Dutch to run the study sessions in Dutch). Additionally, recruits needed to pass the following exclusion criteria: participants needed to be between the ages of 18 and 65 (for sufficient maturity, possible health concerns, and technical proficiency), should not be particularly susceptible to motion sickness (this is related to cybersickness), and could not have significant hearing or vision impairments.

Groups of three were chosen for the more complex interactions that may occur that in groups of two. These sets may have different interactions when more uncertainty is at play regarding communication. Despite this preference, sessions would be run with groups of two if the third participant did not appear.

In total, 29 recruits were gathered from TNO Soesterberg's Internal Recruitment System. Additionally, a set of TNO Interns with no or limited knowledge on the study were on standby to fill in for absences, if they occurred. Despite 29 slots being filled, only 19 total recruits and fill ins across 8 groups actually participated. Due to absences, cancelled timeslots, and people showing up at the wrong times, it became necessary to run the experiment with groups of two participants for some sessions instead of the preferred 3 participants. 5 groups were run with two participants and 3 groups were run with three participants. Of the participants, 13 were female and 6 were male. The average age of participants was 39 (*M=39.05, SD=14.71*).

4.3 Protocol

The experiment was conducted according to the following protocol:

1. Participants arrive and introduced to each other.
    a. If applicable, backup participant placed. (5min)
2. Participants given opportunity to use restroom facilities.
3. Participants taken to the Immersive Collaboration Lab (ICL). (3min)
4. Participants given the information letter and consent forms to fill out. (5min)
5. Participants given Introduction sheet to read through, which contains the task they need to perform. (3min)
6. Participants perform height, controller, and Inter Pupillary Distance (IPD) calibration for visual clarity while using the headset. (3min)
7. Participants familiarize with themselves in an isolated virtual reality environment. (3min)
8. Participants perform the Task in the first room of the first set. (7min)
9. Participants fill out Ad-Hoc questionnaire about Room details. (3min)
10. Participants perform the Task in the second room of the first set. (7min)
11. Participants fill full set of questionnaires, Avatar 2 prepared. (10min) - All Self-Reported Dependent Variables Measured in the Questionnaire.
12. Participants familiarize with themselves in an isolated virtual reality environment. (3min)
13. Participants reminded of task (come up with changes they would make). (1min)
14. Participants perform the Task in the first room of the second set. (7min)
15. Participants fill out Ad-Hoc questionnaire about Room details. (3min)
16. Participants perform the Task in the second room of the second set. (7min)
17. Participants fill questionnaires. (10min) - All Self-Reported Dependent Variables Measured in the Questionnaire.
18. Participants asked how they feel the experiment went. (5min)
19. Participants thanked, then escorted out.

Task Protocol:

1. Participants enter room environment; technical check. (1min)
2. Participants travel through room environment and discuss it amongst themselves. 5min)
3. Participants leave VR. (1min)

4.4 Survey

In order to adequately gather quantitative information about different layers of the brainstorming and evaluation process, different variables would need to be represented. Gathering information about the individual rooms and group discussions is best suited after each room. This keeps the time between the workspace activity and the response low, which will mitigate issues related to recall and recency. As such, a short survey comprised of ad-hoc questions related to the most recent room and most recent discussion will be administered after the first room in each room set. These ad-hoc questions will contain several Likert Scale questions that relate to how well the participant liked the most recent room. In order to get qualitative information and see how well the participants' responses align with one another, open ended questions listing a participant's perceived Pros, Cons, and Recommendations regarding the most recent room experienced will also be present.

Several excerpts from validated questionnaires would be used to gather information on other related variables. These questionnaires will be 7-point Likert Scales and may be slightly reformatted or rephrased to be presented in a consistent way with each other. However, the blend of these questionnaires would be administered as a longer survey after the second room in each room set, giving participants enough time and experience to answer the questions that would better reflect the avatars impacts with less interference from any individual room.

Appendix B below contains the full list of variables examined that originate from another source, which source they were from, and structural changes to the questions that occurred. Additionally, Appendix C contains the list of Ad Hoc questions asked in the short and long surveys. Appendices F and G contains the short and long surveys given, respectively.

4.5 Environments

The environments used would need to be workspaces everyone was familiar with. While offices are often the first thing that comes to mind when working, they present several problems. Offices are not uniform in their features and often have significant differences based on profession, organization values, and amenities on-site. This gives offices a much broader range of possibilities, which makes it harder to comment on them if you are not in the profession that it is built for. Offices are also problematic when considering stay-at-home parents, the unemployed, those who primarily or exclusively work remotely, or a number of other demographics which do not experience regular trips to a location.

However, the study calls for a need for evaluations in a setting with workplaces. Due to the broad demographics used for testing, this needs to be a workplace setting that everyone has been in and has some ability to comment on. With this in mind, kitchens were selected as a functional and familiar environment for the study. Kitchens are dedicated spaces broken up into stations with dedicated equipment and are in practically every home. People may also be peculiar about how their kitchen is set up and may have requirements that easily come to mind during discussion. Furthermore, kitchens are workplaces that often include multiple people working in the same space simultaneously, making it sensible to evaluate the design in small groups. In order to test across different kitchens, each group would go through two kitchen environments while in a given state of fidelity. These kitchens would be in a similar, moderate level of fidelity, with the models having no windows, space to walk around, and a number of appliances throughout the kitchen environment. The kitchens used are available in Appendix D.

4.6 Avatars

The avatars in different states of fidelities are at the core of this study. In order to have a significant difference in avatar states, the high fidelity avatars would have a higher degree of realism and the low fidelity avatars would need to be more abstract. When using a more realistic model, full body avatars would be more preferable than avatars representing only a portion of the body due to the implications avatar completeness may have on fidelity. This requirement for completeness necessitates use of inverse kinematics on high fidelity avatars.

Each set of avatars needed to have approximately the same functionalities, but demonstrated in different ways. This requirement meant that even though the high fidelity avatars were going to be more realistic, any additional features for realism needed to be present in the lower fidelity avatars. Due to its importance in communication and reducing uncanniness, eye tracking would need to be added to both models.

Table 2

Avatar Resources Considered

Name (Source)	Full Body	User Customization	Unity Integration	Headset Compatible	Fidelity	Tracking Modalities
Ready Player Me (Tõke et al., 2024)	Yes	Yes	Yes	Oculus	Low	Hand, Lip-sync
Meta Avatars (Meta, 2024)	No	Yes	Yes	Oculus	Low	Eye, Hand, Body
MetaHuman Creator (Epic Games, 2024)	Yes	Yes	No	Vive, Oculus	High, Adjustable	Eye, face, head, body
Rocketbox (Gonzalez-Franco et al., 2020)	Yes	Yes (Manual)	Yes	Vive, Oculus	High, Adjustable	Eye, face, head
Mixamo (Corazza & Kareemi, 2024)	Yes	No	Yes	Vive, Oculus	Avatar Dependent	Body
Viverse (HTC Corporation, 2024)	Yes	Yes (Mobile App)	Yes (UniVRM)	Vive	High	Eye, Lip-sync, Face, Hand

While the Low Fidelity Avatars are more abstract and comprised of simple shapes and simple textures, High Fidelity Avatars are complex. For the purposes of this study, use of externally constructed High Fidelity Avatars was most practical. A fair number of avatars are available, though consistent styling and consistent formatting is important to ensuring that it development is consistent and maintain no conflicts with the visual quality and tone of each avatar. Additionally, they need to be compatible with the aforementioned requirements. Due to this, Microsoft Rocketbox was chosen as the right mix of high visual fidelity, full body format, and technical compatibility.

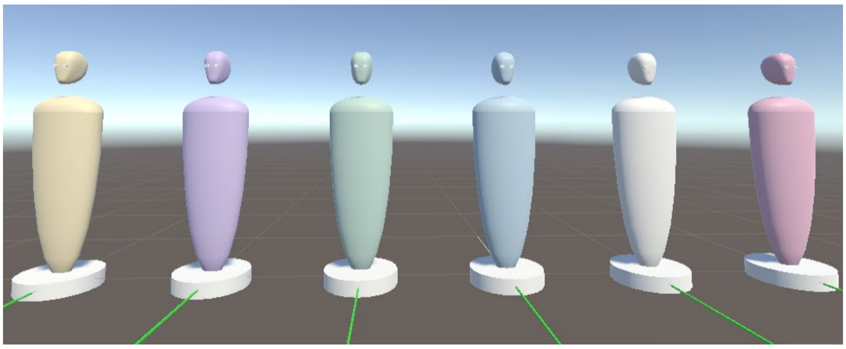


Figure 1: Abstract Avatars in Pastel Colors (Hands Not Visible)

Since we wanted users to select an avatar that felt most like them, choosing avatars was something considered to be necessary, especially for the more realistic avatars. Six male and six female Rocketbox avatars were selected as potential participant avatars (Figure 2). In order to balance this out and make sure that the Low Fidelity Avatars were distinct, six Low Fidelity Avatars in different pastel colors were offered to chose from (Figure 1). In order to avoid confusion, duplicate avatars were not permitted.

Figure 2: Selected Microsoft Rocketbox Avatars

4.7 Data Analysis

Following completion of data collection, both qualitative and quantitative results were processed and analyzed. Quantitative data analysis was accomplished using Python with the Pandas, NumPy, and SciPy libraries and data stored in CSV files. After checking for normality, quantitative data between differences in avatar fidelity groups were tested with paired *t*-tests. Following this, additional post-hoc tests relating avatar fidelity and environments were performed.

Additionally, qualitative analyses were performed with information collected from observations and open-ended survey questions. The raw data was sorted into different codes, which were then recategorized. Following this, patterns in frequency of related codes relating to different conditions and circumstances were recorded.

5 Development

Since the goal of this study was to compare VR brainstorming activities in different states of avatar fidelity, several development requirements became clear. Firstly, Avatars needed to be made in a high and a low state of fidelity. In order to increase realism, steps had to be taken to account for nonverbal cues normally displayed. Due to the role that eye gaze and body movement have on nonverbal communication, eye tracking and inverse kinematics needed to be implemented and a selection of avatars had to be present. In order to make sure that the avatars were not being judged on pure functionality and on fidelity, this meant that eye tracking and avatar selection needed to be implemented in the low fidelity avatars as well.

Doing brainstorming tasks for room design requires rooms to evaluate. Since the environments being evaluated were kitchen prototypes, they needed to be detailed enough to comment on but still relatively simple overall. This meant that the fidelity of these 4 kitchens needed to be moderate, with similar size and features. The lab used had a dedicated VR space, so these kitchens could be fairly spacious.

Additionally, there was a need to coordinate multiple avatars into a single environment that the user could perceive. This meant that networking needed to be implemented in virtual reality. This included a need to coordinate spawning and despawning, and movement between rooms, the actual tracking and movement of the avatars, and information needed for inverse kinematics, eye tracking, and avatar model.

The study was conducted using HTC XR Elite headsets and Vive Full Face Trackers. These headsets have two connected controllers and required calibration to use eye tracking. In order to perform the experiment, avatars and environments needed to be prepared. Building everything entirely from scratch would be unfeasible given the timeframe, so several premade technologies were used. The avatars would need to be usable in the Unity3D environment.

5.1 Inverse Kinematics

Motion for the high fidelity avatars was mapped and coordinated using RootMotion's FinalIK asset (Lang, 2023). This was a fairly robust inverse kinematics system, which had high functionality and many adjustable settings. Each avatar was connected to this, which filled in the movement for body parts not directly gathered by input devices (HMD and hand trackers).

5.2 Eye Tracking

Eye contact and tracking is considered a generally important social cue, so it was considered an important feature to implement. The Vive Full Face Tracker was used to get eye and facial information from each participant while they were in virtual reality. This was used in conjunction with the Unity OpenXR plugin, which allowed this information to be mapped in Unity. In order to test this with fidelity in mind, the same type of functionality needed to be applied to both the high and low fidelity avatars.

5.3 Rocketbox Avatars

The Rocketbox Avatars were constructed using a mesh with many blendshapes present. I made a script that gathered this mesh, the blendshapes, and coordinated them with the eye tracking inputs from the OpenXR Unity plugin. When new visual information was gathered, it updated the avatar accordingly. An example Rocketbox avatar is seen in Figure 3.

Figure 3: Rocketbox (High Fidelity) Avatar Eyes

5.4 Abstract Avatars

Unlike the Rocketbox avatars, the abstract avatars used for the Low Fidelity condition had no such blendshapes could be mapped to. To give an eye shape, oval shaped cylinders acted as the outline for each eye with a square plane inside it (Figure 4). A smaller cylinder was used for the pupil. The blendshape information from the OpenXR plugin was converted into coordinates, which the pupil used to move on the eye plane and emulated eye movement.

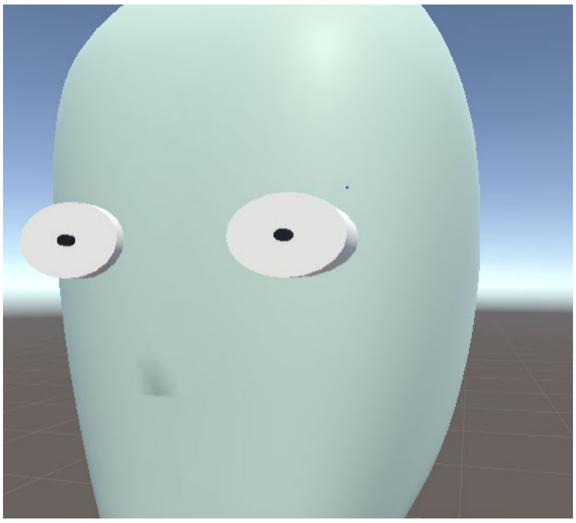


Figure 4: Abstract (Low Fidelity) Avatar Eyes

5.5 Room Setup

The rooms selected were taken as premade models found on SketchUp. In order to have variety, but similar elements, some of the rooms were modified to remove windows and simplify layouts by another TNO employee. Each kitchen had a stove, an island, a sink, cabinets, and some bigger appliances. During a pilot test, color lighting was found to be difficult to navigate due to difficulties seeing the edges of structures in the kitchen without shading, so it was moved to use a gradient lighting scheme, which automatically shaded the environment. Images of these kitchens may be seen in Appendix D.

5.6 Networking

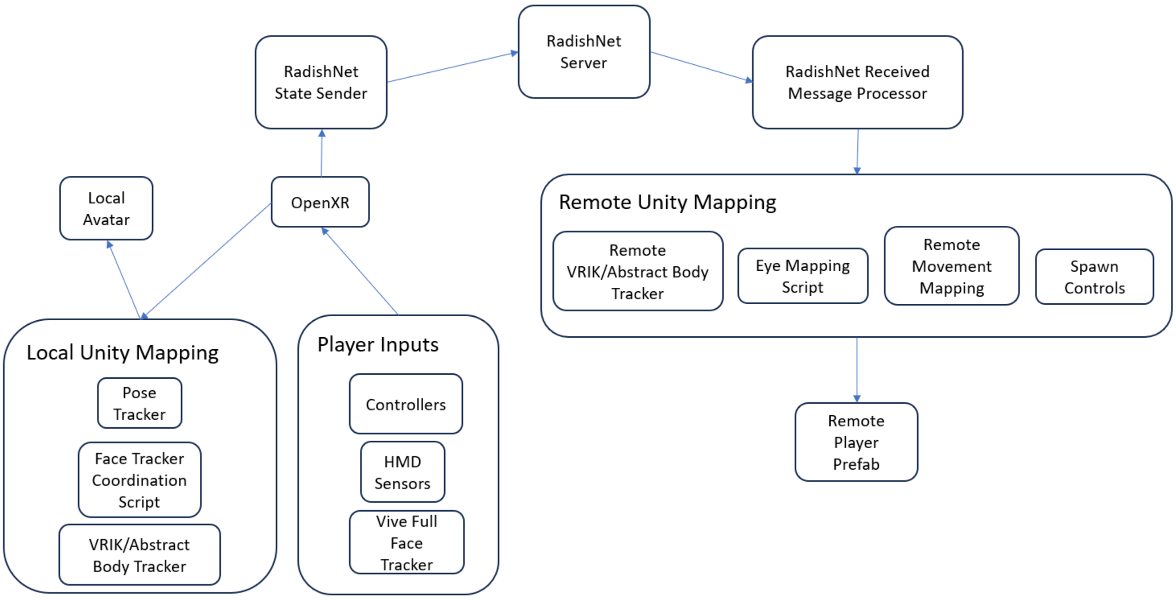


Figure 5: Input Flow Diagram

5.6.1 RadishNet

While participants would be collocated, networking still needed to be implemented to synchronize multiple user's avatars in the same virtual environment. An open-source framework, RadishNet (Wolbers, 2024), was used to coordinate. Radishnet already has a simple and customizable setup, where a server acts as the distributor for information gathered by numerous clients. This made the networking process much easier and shifted the focus towards tailoring the framework to send the correct information and integrate it into the client environments.

For the purpose of this study, this meant gathering each user's model used, avatar head, avatar hands, and eye tracking information and taking the other users' outputs and mapping them to each local avatar representation. When starting the application, an instance of the user's selected avatar would be spawned in each local Unity environment and added to a dictionary of GameObjects (alongside an id). When new information from the state would be sent, the local avatar would be set to match the new information. Additionally, information about the color of the avatar being displayed were also sent. For the low fidelity avatar, the user's bodies were locally attached to follow and rotate according to the location of the user's head. The higher fidelity avatar included mappings to separate FinalIK components to give motion to the Rocketbox avatar's full bodies. The complete map is seen in Figure 5.

5.6.2 Vive Business Streaming

During the study, the Radishnet server coordinated the builds across 3 other computers. While the builds of the Unity environment were capable of being ported to the different headsets, the networking was built to work on computers. The HMDs also depleted battery charge rather quickly, so running additional applications might have strained this more. With this in mind, the Unity builds were run on the computers that had direct, wired networking with the computer acting as a server. The XR Elite HMDs each ran Vive Business Streaming, which connected with the computers over Wi-Fi and streamed the builds.

This had the additional benefit of allowing spawning to be coordinated by keyboard and allowed the researcher to have direct control over which environment the users were located in at a given time. Key inputs would set a new spawn location, which would transfer the participants in the virtual environment upon a second key input.

6 Results

6.1 Avatar Fidelity

During the study, 8 groups were ultimately run with a total of 19 participants. Visual inspection and Shapiro-Wilkes tests were conducted to test for normality when appropriate. A series of paired *t*-tests compared the participants' data based on their current state of avatar fidelity in order to test the hypotheses.

Table 3

Paired t-Test Between Variables in Low and High Fidelity

Variable	RQ	DF	Low Fidelity		High Fidelity		*t*-Statistic	*p*-Value
			M	*SD*	*M*	*SD*		
Co-Presence	A	18	5.703	0.981	5.956	0.989	-1.755	.096
Perceived Message Understanding	A	18	5.921	0.654	5.921	0.874	0.000	1.0
Attention Allocation	A	18	4.982	0.951	5.140	0.961	-0.908	.429
Perceived Affective Understanding	A	18	4.816	1.047	4.974	1.103	-1.302	.209
Spatial Situation Model	B	18	6.036	0.600	6.125	0.589	-1.124	.276
Self-Location	B	18	5.697	0.504	5.855	0.688	-1.294	.212
Satisfaction	C	18	5.877	0.826	5.912	0.736	-0.282	.781
Production Blocking	C	18	6.053	0.705	6.211	0.694	-1.064	.301
Evaluation Apprehension	C	18	5.079	1.250	4.684	1.145	1.204	.244
Free Riding	C	18	6.105	0.737	6.263	0.653	-1.372	.187
Synergy & Stimulation	C	18	5.860	0.697	5.947	0.696	-1.0	.331
Sufficient Time	C	18	5.763	1.475	5.789	1.158	-0.075	.941
Preference Ratings	D	18	3.693	0.925	3.561	0.983	0.507	.618
Number of Pros	D	18	3.026	0.993	3.105	1.209	-0.419	.680
Number of Cons	D	18	3.447	1.657	3.868	1.950	-1.804	.088
Number of Recommendations	D	18	2.895	1.049	2.947	1.383	-0.224	.826

The means, standard deviations, *t*-Statistics, and *p*-Values from the Paired *t*-Tests can be seen in Table 3.

RQ A. Impact on Team Dynamics

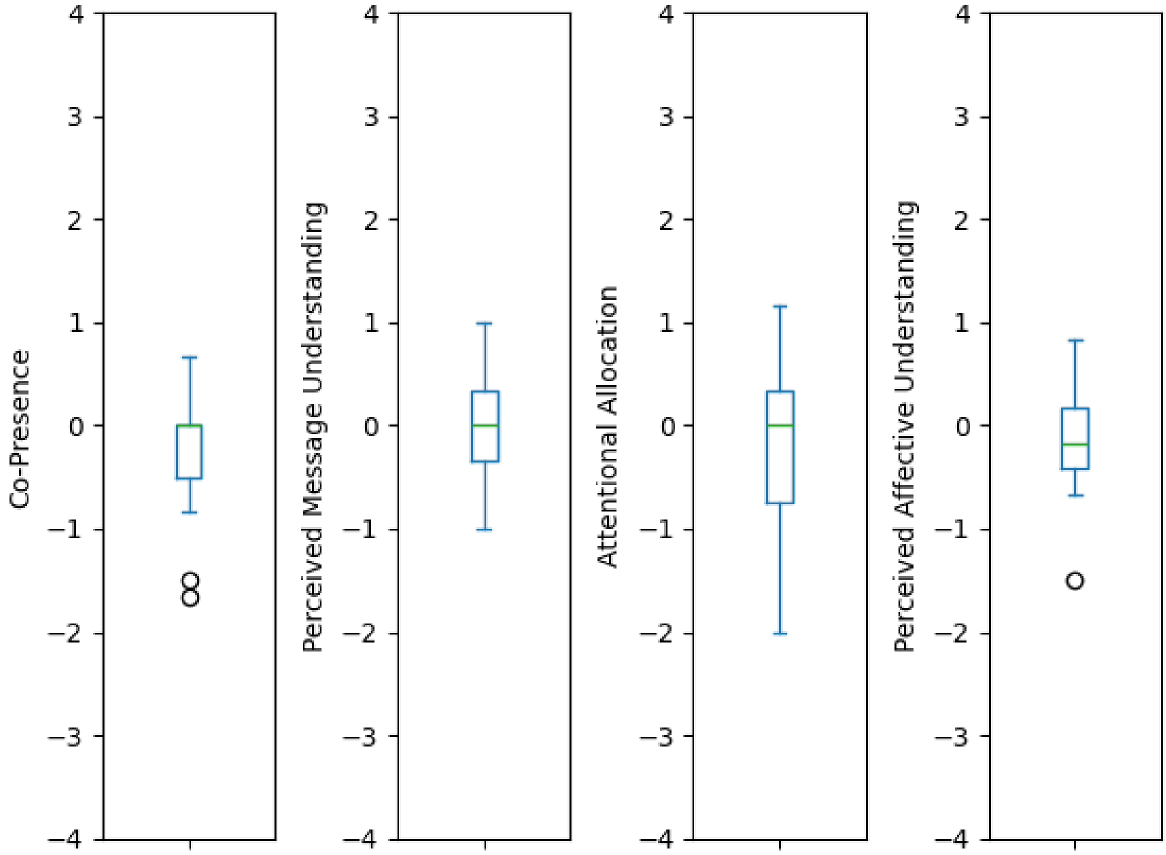


Figure 6: Box plot of difference of means from variables related to team dynamics between Low Fidelity and High Fidelity conditions.

Co-Presence, Perceived Message Understanding, Attentional Allocation, and Perceived Affective Understanding do not have statistical differences between states of Avatar Fidelity, which does not support Hypothesis A. The differences between these scores in High and Low Avatar Fidelity are demonstrated in Figure 6.

RQ B. Influence on Perceptions of the Environment

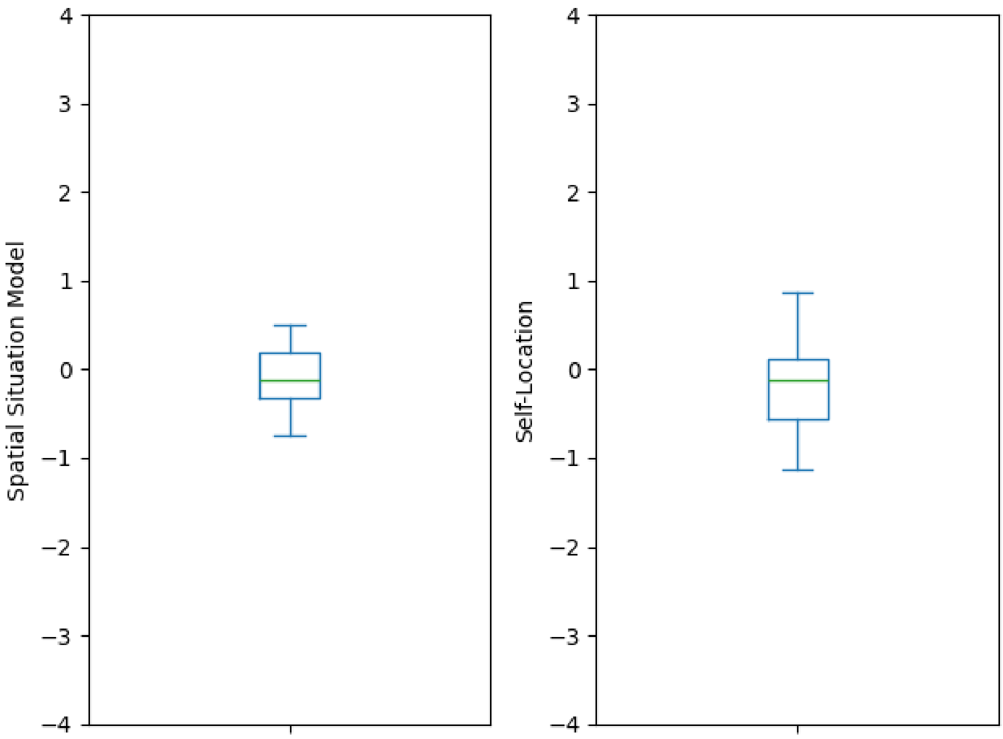


Figure 7: Box plot of difference of means from variables related to perception of the environment between Low Fidelity and High Fidelity conditions.

Spatial Situation Model and Self-Location do not have statistical differences between states of Avatar Fidelity, which does not support Hypothesis B. The differences between these scores in High and Low Avatar Fidelity are demonstrated in Figure 7.

RQ C. Impact of Behaviors While Brainstorming

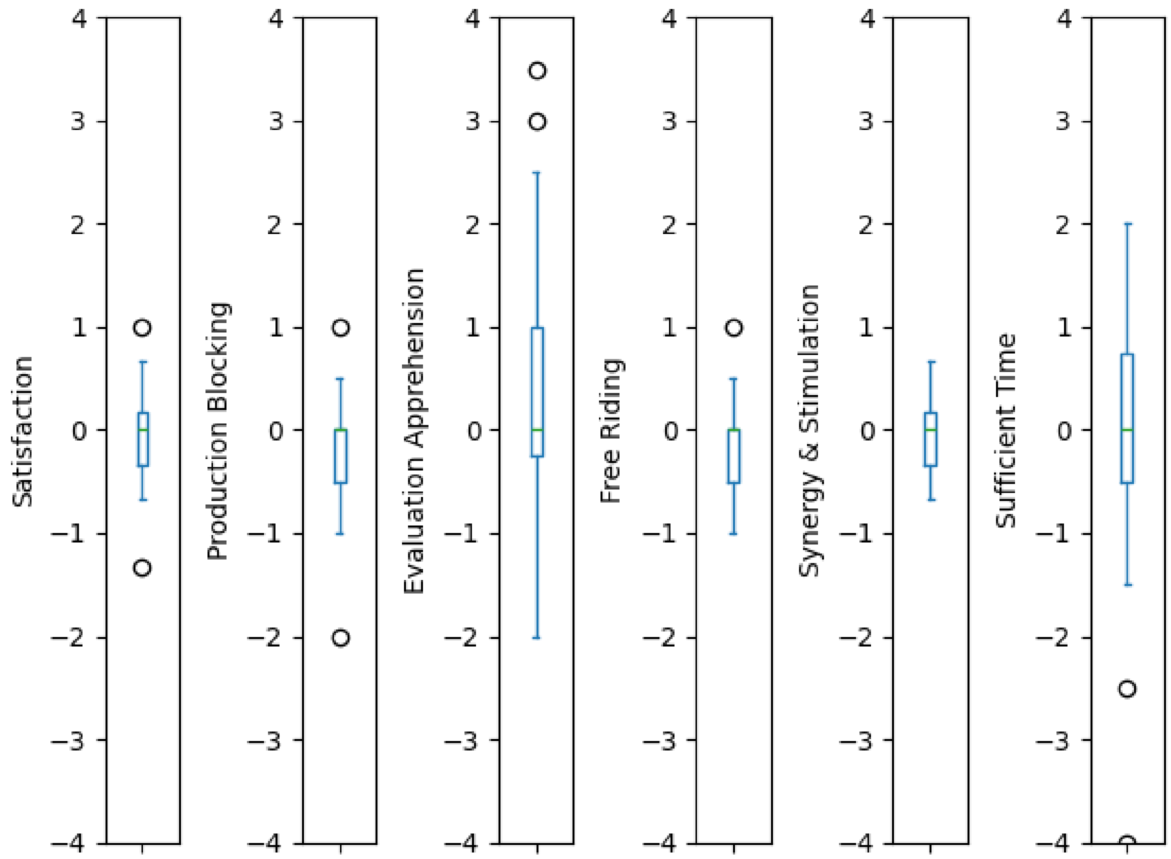


Figure 8: Box plot of difference of means from variables related to behaviors while brainstorming between Low Fidelity and High Fidelity conditions.

Satisfaction, Production Blocking, Evaluation Apprehension, Free Riding, Synergy & Stimulation, and Sufficient Time do not have statistical differences between states of Avatar Fidelity, which does not support Hypothesis C. The differences between these scores in High and Low Avatar Fidelity are demonstrated in Figure 8.

RQ D. Impact on Room Preferences

Preference Ratings, Number of Pros, Number of Cons, and Number of Recommendations do not have statistical differences between states of Avatar Fidelity, which supports Hypothesis D.

6.2 Starting Fidelity

Following the initial results, a series of independent *t*-tests were conducted as an exploratory measure that compared the participants by the fidelity they started with. Using Starting Avatar Fidelity to split the participants into two groups, those that started in High Avatar Fidelity and Low Avatar Fidelity, the data from each room set was examined and there were several significant results discovered. After performing Shapiro-Wilks Tests for normality, *t*-tests were not performed for the number of Pros, Cons, and Recommendations due to some groupings containing results significantly different from normal. Further testing could be performed with Mann-Whitney U Test, but were not due to time constraints. Considering the initial avatars were the ones used while first doing the task and seeing the virtual kitchens, the Avatar Starting Fidelity was checked in order to test if the initial avatars primed participants in a way that framed behavior and perspective even after switching avatars and environments. Hereafter, results associated with each research question and hypothesis will be presented. The results of these *t*-tests regarding Starting Fidelity have been compiled into tables for Room Set A (Table 4) and Room Set B (Table 5).

Table 4

Independent t-Test Between Variables in Low and High Starting Fidelity for Room Set A

Variable	Low Starting Fidelity		High Starting Fidelity		*t*-Statistic	*p*-Value
	M	*SD*	*M*	*SD*		
Co-Presence	6.014	0.747	5.286	1.279	1.580	.132
Perceived Message Understanding	6.042	0.599	5.5	1.076	1.423	.173
Attention Allocation	4.952	0.764	4.959	1.177	-0.0154	.186
Perceived Affective Understanding	5.097	0.777	4.333	1.340	1.587	.131
Spatial Situation Model	6.161	0.570	5.821	0.586	1.239	.232
Self-Location	5.854	0.361	5.554	0.607	1.365	.190
Satisfaction	6.139	0.594	5.429	0.897	2.087	.052
Production Blocking	6.125	0.678	5.786	0.806	0.981	.340
Evaluation Apprehension	6.083	0.669	5.857	1.215	0.528	.604
Free Riding	6.25	0.657	6.0	0.764	0.755	.461
Synergy & Stimulation	5.944	0.372	5.714	0.989	0.734	.473
Sufficient Time	6.541	0.689	5.643	1.345	1.943	.069
Preference Ratings	3.514	1.173	3.5	0.667	0.348	.732

Table 5

Independent t-Test Between Variables in Low and High Starting Fidelity for Room Set B

Variable	Low Starting Fidelity		High Starting Fidelity		*t*-Statistic	*p*-Value
	M	*SD*	*M*	*SD*		
Co-Presence	6.347	0.505	5.167	1.155	3.114	.006
Perceived Message Understanding	6.167	0.663	5.174	0.737	1.378	.186
Attention Allocation	4.857	0.723	4.735	1.359	0.259	.800
Perceived Affective Understanding	5.347	0.773	4.333	1.323	2.127	.048
Spatial Situation Model	6.302	0.537	5.821	0.633	1.765	.096
Self-Location	6.031	0.693	5.429	0.624	1.892	.076
Satisfaction	6.194	0.460	5.429	1.013	2.280	.036
Production Blocking	6.458	0.498	5.929	0.787	1.809	.088
Evaluation Apprehension	6.0	0.739	5.714	1.604	0.535	.600
Free Riding	6.417	0.557	5.857	0.852	1.740	.100
Synergy & Stimulation	6.083	0.452	5.714	1.079	1.053	.307
Sufficient Time	5.875	1.090	4.429	1.539	2.401	.028
Preference Ratings	3.514	0.805	4.143	0.979	0.252	.804

RQ A. Impact on Team Dynamics

There was a significant difference in Co-Presence and Perceived Affective Understanding between Starting Avatar Fidelity conditions for Room Set B.

There was not a significant difference in Co-Presence, Perceived Message Understanding, Attentional Allocation, or Perceived Affective Understanding in Room Set A.

There was not a significant difference in Perceived Message Understanding or Attentional Allocation in Room Set B.

RQ B. Influence on Perceptions of the Environment

There was not a significant difference in Spatial Situation Model or Self-Location in Room Set A or Room Set B.

RQ C. Impact of Behaviors While Brainstorming

There was a significant difference in Satisfaction or Sufficient Time between Starting Avatar Fidelity conditions for Room Set B.

There was not a significant difference in Satisfaction, Production Blocking, Evaluation Apprehension, Free Riding, Synergy & Stimulation, or Sufficient Time for Room Set A.

There was not a significant difference in Production Blocking, Evaluation Apprehension, Free Riding, or Synergy & Stimulation for Room Set B.

RQ D. Impact on Room Preferences

There was not a significant difference in Preference Ratings between Starting Avatar Fidelity conditions in Room Set A or Room Set B.

6.3 Qualitative

There were 25 mentions of stylization across the different rooms when in a state of High Avatar Fidelity and only 10 while in a state of Low Avatar Fidelity. These mentions included perceived Pros, Cons, and Recommendations regarding the rooms stylization. Some included terms about the room being "modern", "old-fashioned", "industrial", and "luxurious" (amongst other things).

In different states of initial avatar fidelity, there were different ways that the environment was interpreted. The participants that started in a state of High Fidelity noted 27 things missing across the rooms while those that started in a state of Low Fidelity found 10 total. Of these, 20 items from the High Starting Fidelity groups and 7 items from the Low Starting Fidelity groups related to appliances and room features that were absent (like power outlets, windows, refrigerators, etc.). These items missing were not necessarily consistent across the different groups, with groups doing things like complimenting the refrigerator or commenting on a missing refrigerator for the same room.

The remaining 7 items from the High Starting Fidelity and 3 items from the Low Starting Fidelity groups were related to missing interactivity in the virtual environment(like opening drawers, cabinets, and changing color settings).

There was limited agreement regarding the Pros, Cons, and Recommendations provided within groups that worked together. In total, of the 32 total opportunities for the different groups to list Pros, Cons, and Recommendations (four rooms and eight groups), there were only 2 times that the majority of Pros provided were listed by all members of a group, 6 times that the majority of Cons provided were listed by all members of a group, and 2 times the majority of Recommendations provided were agreed upon by all members of a group.

6.4 Other Results

Room Order

There were no significant results observed based on the Room Set that a group started with. These results were checked for each variable in each Room Set. This was true for Room Preference scores for each Room Set separately and both combined.

Differences Between Mean Group Scores

There were no significant results observed based on the differences between averaged values for each group.

7 Discussion

7.1 Reflections

7.1.1 Avatar Fidelity

During this study, the quantitative results did not support the idea of significant differences between groups in differing states of Avatar Fidelity. Statistically, there were only marginally significantly results between the tested samples. However, differences between tested groups suggest that higher Avatar Fidelity generally leads to near or improved scores, even if not statistically significant. The absent participants in the experiment may very much have impacted these results, which may have been significant had there been more datapoints. It is worth noting that many paired *t*-tests were performed and not corrected for, so the chance of spurious correlations is possible and future testing is recommended.

Regarding RQ1, Co-Presence, Perceived Message Understanding, Attentional Allocation, and Perceived Affective Understanding are all not significant and therefore do not support the hypothesis. Of these variables, Co-Presence approached significance, which would have been in line with expectations. However, it is plausible that the nature of the tasks at hand would have influenced the amount of time participants spent looking at one another, despite participants occupying the same environment.

The variables used for RQ2 also were largely not significant. Both Spatial Situation Model and Self-Location had non-significant results. This was not expected, given the differences in models. I would have expected this to be significant and given the size of the environments and differences in physical traits of the avatars. Further, I would have suspected that desynchronization would have lowered scores. However, both Spatial Situation Model and Self-Location were rated highly in both conditions, which indicates that participants were able to process the environment and understood their location within it well.

No activities relating to brainstorming were significantly different from one another. This leans against the hypothesis for RQ3, which estimated that there would be differences in behavior. This might further have been impacted by the absent participant, given that the counterbalancing between rooms was impacted (Appendix A).

There were no significant differences between preferences scores between Room Set A and Room Set B. This supports the hypothesis for RQ4. Interestingly, there was comparatively little consensus on listed Pros, Cons, and Recommendations within a given group. This might have been due to a lack of outside structure, but it also may have been standard. Regardless, no differences were found.

7.1.2 Influence of Starting Fidelity

While exploratory in nature, the starting level of avatar fidelity was found to have numerous significant effects. The extent of starting fidelity was not even across both Room Sets, but was found to be associated with statistically significant differences in both. This is interesting, because it is evidence towards the avatar fidelity having a clear influence on the way the users interact in VR brainstorming activities.

It is not completely unexpected that there was some impact with this. Priming through avatars has been shown to influence behavior via the Proteus Effect (Peña et al., 2009). Additionally, the influence of the first avatar of a series of avatars having an impact on subsequent perspectives was observed before (Ma & Pan, 2022). This is interesting, as there may be functional applications to different avatars in activities where an avatar is not necessarily the focus. There were no mirrors to look at nor did the activity require participants to stare at one another, but still influenced the outcomes of the activity. It is interesting though that this influence appeared to be more significant on framing perspectives than the avatars that were used in a given moment. Perhaps the initial state of avatar fidelity influences the formation of mental models for the rooms, activities or groups that are being performed and persists as it gets built upon.

Despite the primary observation being statistically insignificant, this leans towards moderately supporting Hypothesis A. While only significant in Room Set B, Co-Presence and Perceived Affective Understanding were both significantly different depending on starting fidelity. However, both Attentional Allocation and Perceived Message Understanding were not found to have significant differences based on Avatar Fidelity in either Room Set. It is interesting that these effects were not seen in the current state of Avatar Fidelity and were only significantly aligned with the Starting Avatar Fidelity. This leads me to suspect that initial degree Avatar Fidelity is part of the equation of impacting team dynamics, but is part of a broader sense of how environmental tone impacts communication.

7.1.3 Perceptions of Environment

Neither the current state of Avatar Fidelity nor Starting Avatar Fidelity demonstrate significant differences regarding the quantitative data collected. The quantitative evidence gathered does not support Hypothesis B. However, the qualitative data does show some interesting differences in how the participants interacted with the environment and the expectations that they have.

7.1.3.1 Stylization

The participants that started in a state of High Avatar Fidelity were much more likely to make comments about the stylization of the environment they were in than those that started in a state of Low Avatar Fidelity. While its possible that it was coincidental, the difference in commentary related to stylization is rather expansive despite all participants going through the same rooms. “Kitchens are not clean, so it was uncomfortable to see one so clean. It felt sterile and unnatural.” (Participant 18). The state of cleanliness and sterility of the environments was something viewed as a negative, while modern stylization was generally viewed in a more positive manner. This might say something about comparison to real world environments, where realistic avatars may have led to expectations about a more visually realistic environment.

7.1.3.2 Desire for Interactivity

However, there being possible impacts on the outcomes of collaborative, creative activities is interesting when considering the practical implications of this. One such impact was noticed in the collected qualitative data, where participants that started in a state of Low Avatar Fidelity noticeably were more interested in extra functionalities for interacting with the environment. Despite a higher fidelity avatar being more realistic, the initial lower fidelity may have had some sort of influence on the way the participants fit within the environment. "The realistic avatars felt more like reality, the abstract ones felt like being in a game." (Participant 14). Some participants commented on this, saying that being in the Low Avatar Fidelity felt more like a game, though there was disagreement if it was more or less enjoyable than being in High Avatar Fidelity.

7.1.4 Impact of Behaviors While Brainstorming

7.1.4.1 Lack of Consensus

Something surprising was the generally low level of consensus between participants on the Pros, Cons, and Recommendations for the environments observed within members of the same group. Unique outputs that were not mentioned by other participants were fairly commonplace and there were rarely cases that there was full agreement for the majority of Pros, Cons, and Recommendations listed. While they rarely contradicted each other, different participants in the same group did not list the same things as one another in most cases. Generally, there were a couple key items listed that achieved consensus, but several others that did not.

This is something with many possible explanations applicable. Perhaps this was pre-baked into the circumstances, as kitchens can be personal and may be dependent on external needs; a person with a large family and many guests may have very different circumstances they are considering when compared to somebody living on their own. While discussing with each other, participants may have paid attention to topics prioritized and neglected other comments with little discussion, which left harder to convey concerns and suggestions to the side. This might be distinct from Production Blocking or Evaluation Apprehension and focus more on clarity and ability to articulate a thought. Another possibility is that participants spent most of their time going over a similar set of details, but did not bring up new ones. Alternatively, some things written down may have been thought of after the brainstorming activity instead of during it. Regardless of the reason, there was a lack of alignment between participants and their opinions, which means that the room prototype and activity in its current form did not lean to a mutual sense of the rooms and their features.

7.1.4.2 Awareness of Eye Tracking

When interviewing after the study, very few participants noticed eye tracking explicitly and none brought it up before I mentioned it. The simplest reason I could think of for this was that they simply did not spend a lot of time looking each other in the eyes while focusing on the environment. It could also be that they did not move their eyes enough to notice the difference in most cases. It is worth noting that VR headsets still have limited peripheral vision, so this may have also narrowed the scope of view to the point where they did not notice glancing as much. Even so, it is still undetermined whether or not it was impactful in facilitating social interaction. There was a difference in Perceived Affective Understanding while in Room Set B, which suggests that the avatars led to some difference in emotional conveyance. However, this was not the case in Room Set A, which suggests that this difference is either circumstantial or there are confounding factors at play that are not being observed in reference to this.

7.1.4.3 Sufficient Time

Interestingly, Sufficient Time was statistically different when looking at Starting Avatar Fidelity in Room Set B. While the rating for sense of Sufficient Time was still satisfactory even in the High Fidelity groups, there may be reasons for this. Considering the number of significant differences, it is possible that Room Set B had more features to comment on than the other rooms, compressing time. However, it is also plausible that time pressure was strained by gaps in Perceived Affective Understanding and Co-Presence. The number of characters regarding Room Set B and Room Set A were not significantly different from one another, which supports the idea that there was not necessarily a difference in raw bulk output. However, it is also plausible that this is a constraint of the time given and that there might have been a difference if each group had more time to discuss with one another.

7.1.5 Room Preferences

The absence of significant quantitative differences in room preferences is interesting, given the diversity of observations across each room. It is possible that the reasons for this are the comparatively similar virtual environments. There were not strong disagreements observed through the course of the study and participants generally considered the experience enjoyable in discussion afterwards. While there were not a lot of differences in preferences, the rooms were fairly consistently viewed in a low light. However, it is also worth noting that participants were observed comparing rooms to each other in the different room sets. While doing this, they discussed features that they liked and were present in one room or the other. These reference points may point to problems noticed in one environment that were present in another, but may not have been noticed beforehand. This may have been a results of the activity at play- where participants just discussed the negatives of the rooms and recommendations on how to improve them. Even despite this benchmarking strategy, the rooms were considered similarly bad in each set of rooms.

7.1.6 Room Order

There were no significant results observed based on the Room Set that a group started with.

7.2 Limitations

7.2.1 Language Spoken

The participants in the study all spoke English as a second language, which may have influenced some of their word choice in qualitative evaluation and the way they interpreted some of the questions. In particular, they struggled with the word 'apprehension', which was relevant to a question in the study. This had led to a large variety of words and phrases being used, and sometimes Dutch words were used to substitute for English ones until a substitute was found.

Furthermore, some participants may have overestimated English proficiency. While the vast majority of participants demonstrated strong English skills, some of the participants struggled to comprehend instructions, especially through menus. In these situations, these participants would shift conversations in English to Dutch. This may have also impacted the ability to express the insights from discussions in English, which was required for the study.

7.2.2 Absent Participants

Of the 30 open slots and 29 recruits, only 19 people who agreed to participate appeared in the study. This discrepancy led to only eight of the ten planned groups to run. Of these, only 3 groups had the full 3 people and the remainder were run with 2. This weakened the dataset, added extra inconsistencies between groups, and caused participants who showed up when there were many absents to have to go home without participating in the study. There was likely an impact, because the study was designed to run with 3 participants, with the capacity to run with 2 participants being a backup plan. Unfortunately, the rate of cancelation or absenteeism for the study was comparatively high and groups with two participants were the majority of cases.

7.2.3 Kinematics Limitations

While the Inverse Kinematics added to the avatars, issues with them were generally more distracting. When controllers were not detected by the headset, limbs would stay in place while moving. Occasionally, limbs clipping through the environment and bodies were distracting at times. Further still, changes to height or perspective changed the height of the avatars- while height was set at the start of the study, it could be impacted if the menu control buttons were held or if the floor was not set to the group properly at the start. Additionally, the wrists were not very capable of twisting in a believable way, so they were locked in place. This reduced the natural range of movement that the hands could do, which could have impacted things.

7.2.4 Technical Proficiency

Due to the language and experience barrier experienced by some of the older participants, they did not have as strong of technical proficiency and it made it difficult to give and sure instructions were being followed. While this was anticipated beforehand and demonstrations on how to use the VR controllers were given beforehand, some participants still struggled to follow. This led to occasional incidents where the older participants would select or hold the menu button while trying to select menu options (which was necessary to calibrate the Vive Full Face Tracker and set floor position).

7.2.5 Self Reporting

Due to time constraints, the majority of quantitatively measured variables were recorded during surveys at the end of every room set. This could have influenced the participant's recall ability. The impacts of this are hard to measure, but it may have limited results- especially when considering the shift between avatars occurred right before the start of each room set. This means that differences in adaption to the avatars might not have been measured fully.

7.2.6 Room Sets

The time constraints also caused a limited distribution of rooms. While starting avatar fidelity and starting room set were counterbalanced between the participants, the rooms within each set remained static. This meant there are potentially ordering factors that could have influenced parameters in a way that would be difficult to detect from the dataset. Additionally, the participant deficiency impacted the counterbalancing for the starting room set, which means that there were two more groups that started in Room Set A than Room Set B.

7.2.7 Avatar Selection

While there were a diverse array of avatars available, the tone of the avatars was somewhat inconsistent and options were surprisingly limited in some ways. There were limited numbers of African female avatars outside of specialized jobs. Furthermore, the styles of clothing options for the standard avatars were not consistent. There was an effort made to choose avatars with relatively neutral stylization (not religious or political in nature, not wearing distracting clothing, nor wearing clothing strongly tied with occupational stereotypes), but it was not entirely possible for all demographics. While it was not chosen by any participant, the avatar representing an African female was originally a pilot avatar and still retained bands on the shoulders for the profession. There was no distinct African female in standard attire. While effort was made to mitigate strong differences in avatar tone, it is possible that this could have had some impact on the ways participants saw each other.

7.3 Future Work

7.3.1 Adjusting to Avatar Changes

In this study, participants experienced distinct transitions between avatars. Examining the ways that users take time to adjust to others using different avatars would be useful to understanding the implications of adapting to differences in form and ownership. It would also be interesting to see if there are mitigation tactics. Further, seeing how teammates adapt to changes in avatar might also be interesting, considering that avatars can be viewed as both a representation of someone and part of the environment.

7.3.2 Expectations of Interactivity

While in virtual reality, there was a frequent desire for added interactivity in the environment. This seemed to be related to both the environment and the avatar being used and was focused on the cabinets, refrigerator, and changing the color scheme of the room. This occurred to the more realistic avatars, so it would be interesting to explore how virtual realism generates affordances and signifiers (Norman, 2013) in virtual reality.

7.3.3 Avatar Selection in Groups

Something interesting observed was that people selected avatars of different races and sex than themselves. For example, a frequent avatar pick was an Asian female avatar, but nobody in the study was an Asian female. When asked about this, they stated that they either liked the avatar's outfit or wanted to try being something different. This discrepancy might have also primed behavior in some form or changed the way people viewed each other.

7.3.4 Structure While Brainstorming in VR

There was surprisingly little consensus regarding the kitchen environments. In other design thinking tasks, there is either someone running a workshop or keeping track of time in some capacity. Given that this was fairly freeform, more structure might be effective in establishing more consensus. It might also be interesting to compare this to brainstorming activities that occur in reality.

7.3.5 VR Recommendation Prioritization

The limited overlap between participants' perceived Cons and Recommendations were surprising considering that the scores relating to question about Sufficient Time were generally high. In many cases, one or two items did overlap, but it was generally not the majority of the time. It would be interesting to see if the problems participants agreed on in workspace prototypes were prioritized over other critiques they made. It would also be interesting to see how differences in priority are handled if participants had to create plans on how to deal with the problems that were identified.

7.3.6 Embodied Perspectives in VR Evaluation

While doing the experiment, a participant did not fully set the floor to their feet, which caused them to appear shorter than they were externally. This led to situations where they started considering things as points of interest, such as heights of counters. This also led the group discussing what improvement might be made to make the situation more child friendly. In interview afterwards, the group stated that they found this difference helpful, because it helped them consider things they did not originally find problematic. However, they also found it slightly confusing, due to the difference between where the avatar's head was and where their voice was coming from. If understood, this change in perspective might be capable of uncovering new insights, which might be useful for scenarios where brainstorming and evaluation activities are implemented.

8 Conclusion

This study has investigated the way avatar fidelity influences workspace brainstorming and evaluation activities in virtual reality. Avatar features and an environment were selected and developed in Unity3D. Using both avatars in high and low fidelity, groups of participants went through a series of kitchens and discussed their benefits, problems, and recommendations with one another before answering surveys asking questions about the environment and their opinions on it.

After using paired *t*-tests, a participant's current level of avatar fidelity and the room order were not found to have statistically significant impacts on team dynamics, perceptions of the virtual environment, nor room preferences. This was unexpected given the significant difference between the realistic, high fidelity and abstract, low fidelity avatars. Through exploratory testing with independent *t*-tests, the level of avatar fidelity participants started with was found to have significant impacts in Room Set B, including the variables Co-Presence, Perceived Affective Understanding, Satisfaction, and Sufficient Time.

There were qualitative differences found when using Starting Avatar Fidelity as the variable examined, where those that started in High Fidelity ended up wanting more environmental interactivity and finding the lack of detail in the environment unpleasant. In discussion after the activity, participants generally stated that they enjoyed the activity and found it to be helpful when thinking about creating rooms. It was also found that participants shared only a limited amount of overlap with one another when it came to listing Pros, Cons, and Recommendations individually. Participants were mixed on which avatar fidelity they preferred, with some preferring the simplicity of the low fidelity avatars and other preferring the closer proximity high fidelity avatars have to reality.

Virtual reality has been rapidly developing and there are no signs of this development slowing down. However, multi-user VR is still emerging as a field. Creative uses of exploring simulated 3D spaces are being thought up with all the time, but the functional aspects of this can be complicated when more users are present. This use case might be relevant. This study has demonstrated that while realistic avatars feel more like reality, both high and low fidelity avatars are capable of facilitating meaningful discussion. While simpler, less realistic, and body measurements are harder to contextualize, low fidelity avatars are comparatively easy to make, functional, and may take emphasis away from stylization and more on function while brainstorming. High fidelity avatars are much harder to make, can be uncanny (especially if calibration is off), and might anticipate higher levels of environmental interactivity, but are also may feel more like natural interaction, increase attention to details, and might feel more like reality. Doing evaluation and brainstorming tasks can be complicated. Virtual reality can help make the process easier and the avatars used are not, but the avatars are not the limiting factor of its potential.

References

Abdel Razek, A. R., Van Husen, C., Pallot, M., & Richir, S. (2018). A comparative study on conventional versus immersive service prototyping (VR, AR, MR). *ACM International Conference Proceeding Series*, 1–10. https://doi.org/10.1145/3234253.3234296

Andrei, R., Engermann, O. B., Sørensen, C. G., & Löchtefeld, M. (2023). Examining the Effects of Eye-tracking on Dyadic Conversations in Virtual Reality. *ACM International Conference Proceeding Series*, 392–402. https://doi.org/10.1145/3626705.3627794

Aseeri, S., & Interrante, V. (2021). The Influence of Avatar Representation on Interpersonal Communication in Virtual Social Environments. *IEEE Transactions on Visualization and Computer Graphics*, *27*(5), 2608–2617. https://doi.org/10.1109/TVCG.2021.3067783

Babu, S., Suma, E., Barnes, T., & Hodges, L. F. (2007). Can immersive virtual humans teach social conversational protocols? *Proceedings - IEEE Virtual Reality*, 215–218. https://doi.org/10.1109/VR.2007.352484

Bai, H., Sasikumar, P., Yang, J., & Billinghurst, M. (2020). A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. *Conference on Human Factors in Computing Systems - Proceedings*, 1–13. https://doi.org/10.1145/3313831.3376550

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of Social Behavior : Direct Effects of Trait Construct and Stereotype Activation on Action. *Journal of Personality and Social Psychology*, *71*(2), 230–244.

Baron, R. A. (1988). Negative Effects of Destructive Criticism: Impact on Conflict, Self-Efficacy, and Task Performance. *Journal of Applied Psychology*, *73*(2), 199–207. https://doi.org/10.1037/0021-9010.73.2.199

Basori, A. H., & Ali, I. R. (2013). Emotion Expression of Avatar through Eye Behaviors, Lip Synchronization and MPEG4 in Virtual Reality based on Xface Toolkit: Present and Future. *Procedia - Social and Behavioral Sciences*, *97*, 700–706. https://doi.org/10.1016/j.sbspro.2013.10.290

Boletsis, C., Karahasanovic, A., & Fjuk, A. (2017). Virtual bodystorming: Utilizing virtual reality for prototyping in service design. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10324 LNCS*(i), 279–288. https://doi.org/10.1007/978-3-319-60922-5_22

Bovo, R., Giunchi, D., Alebri, M., Steed, A., Costanza, E., & Heinis, T. (2022). Cone of Vision as a Behavioural Cue for VR Collaboration. *Proceedings of the ACM on Human-Computer Interaction*, *6*(CSCW2). https://doi.org/10.1145/3555615

Butts, C. T. (2008). A Relational Event Framework for Social Action. *Sociological Methodology*, *38*(1), 155–200.

Camacho, L. M., & Paulus, P. B. (1995). The Role of Social Anxiousness in Group Brainstorming. *Journal of Personality and Social Psychology*, *68*(6), 1071–1080. https://doi.org/10.1037/0022-3514.68.6.1071

Casanueva, J., & Blake, E. (2001). The Effects of Group Collaboration on Presence in a Collaborative Virtual Environment. *Proceedings of the 6th Eurographics Workshop on Virtual Environments, EGVE 2000*, *January 2001*. https://doi.org/10.1007/978-3-7091-6785-4_10

Cassar, C., Simpson, R. J., Bradbeer, N., & Thomas, G. (2019). Integrating virtual reality software into the early stages of ship design. *RINA, Royal Institution of Naval Architects - 19th International*

Conference on Computer Applications in Shipbuilding, ICCAS 2019, *1*(September). https://doi.org/10.3940/rina.iccas.2019.1.4

Corazza, S., & Kareemi, N. (2024). *mixamo*. Adobe. https://www.mixamo.com/#/

Curhan, J. R., Labuzova, T., & Mehta, A. (2021). Cooperative Criticism: When Criticism Enhances Creativity in Brainstorming and Negotiation. *Organization Science*, *32*(5), 1256–1272. https://doi.org/10.2139/ssrn.3700258

Dennis, A. R., & Valacich, J. S. (1993). *Computer Brainstorms: More Heads Are Better Than One* (pp. 531–537). American Psychological Association.

Diehl, M., & Stroebe, W. (1987). Productivity Loss In Brainstorming Groups : Toward the Solution of a Riddle. *Journal of Personality and Social Psychology*, *53*(3), 497–509.

Epic Games. (2024). *MetaHuman Creator*. Epic Games. https://www.unrealengine.com/en-US/metahuman

Forens, M., Bonnardel, N., & Barbier, M. L. (2015). How communication modalities can impact group creativity in multi-user virtual environments. *ACM International Conference Proceeding Series*, *01-03-Jul-*, 3–6. https://doi.org/10.1145/2788412.2788439

Freeman, G., Acena, D., McNeese, N. J., & Schulenberg, K. (2022). Working Together Apart through Embodiment. *Proceedings of the ACM on Human-Computer Interaction*, *6*(GROUP), 1–25. https://doi.org/10.1145/3492836

Gallupe, R. B., Dennis, A. R., & Cooper, W. H. (1992). *Electronic Brainstorming and Group Size*. *35*(2), 350–369.

Gonzalez-Franco, M., Ofek, E., Pan, Y., Antley, A., Steed, A., Spanlang, B., Maselli, A., Banakou, D., Pelechano, N., Orts-Escolano, S., Orvalho, V., Trutoiu, L., Wojcik, M., Sanchez-Vives, M. V., Bailenson, J., Slater, M., & Lanier, J. (2020). The Rocketbox Library and the Utility of Freely Available Rigged Avatars. *Frontiers in Virtual Reality*, *1*(November), 1–23. https://doi.org/10.3389/frvir.2020.561558

Guegan, J., Nelson, J., & Lubart, T. (2017). The Relationship between Contextual Cues in Virtual Environments and Creative Processes. *Cyberpsychology, Behavior, and Social Networking*, *20*(3), 202–206. https://doi.org/10.1089/cyber.2016.0503

Hammick, J. K., & Lee, M. J. (2014). Do shy people feel less communication apprehension online? the effects of virtual reality on the relationship between personality characteristics and communication outcomes. *Computers in Human Behavior*, *33*, 302–310. https://doi.org/10.1016/j.chb.2013.01.046

Han, E., Miller, M. R., DeVeaux, C., Jun, H., Nowak, K. L., Hancock, J. T., Ram, N., & Bailenson, J. N. (2023). People, places, and time: a large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication*, *28*(2). https://doi.org/10.1093/jcmc/zmac031

Han, E., Miller, M. R., Ram, N., Nowak, K., & Bailenson, J. N. (2022). Understanding Group Behavior in Virtual Reality: A Large-Scale, Longitudinal Study in the Metaverse. *72nd Annual International Communication Association Conference*, 1–27.

Harms, C., & Biocca, F. (2004). Internal Consistency and Reliability of the Networked Minds Measure of Social Presence. *Seventh Annual International Workshop: Presence 2004*, 246–251.

Heidicker, P., Langbehn, E., & Steinicke, F. (2017). Influence of avatar appearance on presence in social VR. *2017 IEEE Symposium on 3D User Interfaces, 3DUI 2017 - Proceedings*, 233–234.

https://doi.org/10.1109/3DUI.2017.7893357

Higgins, D., Egan, D., Fribourg, R., Cowan, B., & McDonnell, R. (2021). Ascending from the valley: Can state-of-the-art photorealism avoid the uncanny? *Proceedings - SAP 2021: ACM Symposium on Applied Perception*. https://doi.org/10.1145/3474451.3476242

Higgins, D., Zhan, Y., Cowan, B. R., & McDonnell, R. (2023). Investigating the effect of visual realism on empathic responses to emotionally expressive virtual humans. *Proceedings - SAP 2023: ACM Symposium on Applied Perception*. https://doi.org/10.1145/3605495.3605799

HTC Corporation. (2024). *Viverse*. HTC Corporation. https://www.viverse.com/

Jenks, C. J. (2009). When is it appropriate to talk ? Managing overlapping talk in multi-participant voice-based chat rooms. *Computer Assisted Language Learning*, *22*(1), 19–30. https://doi.org/10.1080/09588220802613781

Jetter, H. C., Rädle, R., Feuchtner, T., Anthes, C., Friedl, J., & Klokmose, C. N. (2020). "in VR, everything is possible!": Sketching and Simulating Spatially-Aware Interactive Spaces in Virtual Reality. *Conference on Human Factors in Computing Systems - Proceedings*, 1–16. https://doi.org/10.1145/3313831.3376652

Jokinen, K., Furukawa, H., Nishida, M., & Yamamoto, S. (2013). Gaze and turn-taking behavior in casual conversational interactions. *ACM Transactions on Interactive Intelligent Systems*, *3*(2). https://doi.org/10.1145/2499474.2499481

Jovanovic, N., Op Den Akker, R., & Nijholt, A. (2006). Addressee identification in face-to-face meetings. *EACL 2006 - 11th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Conference*, 169–176.

Khojasteh, N., & Won, A. S. (2021). Working Together on Diverse Tasks: A Longitudinal Study on Individual Workload, Presence and Emotional Recognition in Collaborative Virtual Environments. *Frontiers in Virtual Reality*, *2*(June), 1–24. https://doi.org/10.3389/frvir.2021.643331

Kocur, M., Schauhuber, P., Schwind, V., Wolff, C., & Henze, N. (2020). The Effects of Self- And External Perception of Avatars on Cognitive Task Performance in Virtual Reality. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*. https://doi.org/10.1145/3385956.3418969

Köping Olsson, B., & Florin, U. (2011). Idea Exchange and Shared Understanding: Tools Stimulating Thought and Conveying Ideas Idea Exchange and Shared Understanding: Tools Stimulating Thought and Conveying Ideas. *DESIGN PRINCIPLES AND PRACTICES: AN INTERNATIONAL JOURNAL*, *5*. http://www.design-journal.comhttp//www.commongroundpublishing.com/software/

Krum, D. M., Kang, S. H., Phan, T., Dukes, L. C., & Bolas, M. (2016). Head mounted projection for enhanced gaze in social interactions. *Proceedings - IEEE Virtual Reality*, *2016-July*, 209–210. https://doi.org/10.1109/VR.2016.7504727

Lang, P. (2023). *Final IK* (2.3). Root Motion. http://root-motion.com/

Latoschik, M. E., Roth, D., Gall, D., Achenbach, J., Waltemate, T., & Botsch, M. (2017). The effect of avatar realism in immersive social virtual realities. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, *Part F1319*. https://doi.org/10.1145/3139131.3139156

Liao, J., & Wang, H. C. (2019). Gestures as intrinsic creativity support: Understanding the usage and function of hand gestures in computer-mediated group brainstorming. *Proceedings of the ACM on Human-Computer Interaction*, *3*(GROUP). https://doi.org/10.1145/3361124

Ma, F., & Pan, X. (2022). Visual Fidelity Effects on Expressive Self-avatar in Virtual Reality: First Impressions Matter. *Proceedings - 2022 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2022*, 57–65. https://doi.org/10.1109/VR51125.2022.00023

Maran, T., Furtner, M., Liegl, S., Ravet-Brown, T., Haraped, L., & Sachse, P. (2021). Visual Attention in Real-World Conversation: Gaze Patterns Are Modulated by Communication and Group Size. *Applied Psychology: An International Review*, *70*(4), 1602–1627. https://doi.org/10.1111/apps.12291

Mathis, F., Vaniea, K., & Khamis, M. (2021). Observing Virtual Avatars: The Impact of Avatars' Fidelity on Identifying Interactions. *ACM International Conference Proceeding Series*, 154–164. https://doi.org/10.1145/3464327.3464329

Meta. (2024). *Meta Avatars SDK*. Meta. https://developer.oculus.com/documentation/unity/meta-avatars-overview/

Minas, R. K., Dennis, A. R., & Massey, A. P. (2016). Opening the mind: designing 3d virtual environments to enhance team creativity. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2016-March*, 247–256. https://doi.org/10.1109/HICSS.2016.38

Norman, D. (2013). Fundamental Principles of Interaction. In *The Design of Everyday Things: Revised and Expanded Edition* (Second). Basic Books. https://www.nngroup.com/books/design-everyday-things-revised/

Oh Kruzic, C., Kruzic, D., Herrera, F., & Bailenson, J. (2020). Facial expressions contribute more than body movements to conversational outcomes in avatar-mediated virtual environments. *Scientific Reports*, *10*(1), 1–23. https://doi.org/10.1038/s41598-020-76672-4

Olaosebikan, M., Aranda Barrios, C., Kolawole, B., Cowen, L., & Shaer, O. (2022). Identifying Cognitive and Creative Support Needs for Remote Scientific Collaboration using VR: Practices, Affordances, and Design Implications. *ACM International Conference Proceeding Series*, 97–110. https://doi.org/10.1145/3527927.3532797

Paes, D., & Irizarry, J. (2016). Virtual reality technology applied in the building design process: Considerations on human factors and cognitive processes. *Advances in Intelligent Systems and Computing*, *485*, 3–15. https://doi.org/10.1007/978-3-319-41983-1_1

Peña, J., Hancock, J. T., & Merola, N. A. (2009). The Priming Effects of Avatars in Virtual Settings. *Communication Research*, *36*(6), 838–856. https://doi.org/10.1177/0093650209346802

Praetorius, A. S., Krautmacher, L., Tullius, G., & Curio, C. (2021). User-Avatar Relationships in Various Contexts. *ACM International Conference Proceeding Series*, 275–280. https://doi.org/10.1145/3473856.3474007

Ries, B., Interrante, V., Kaeding, M., & Phillips, L. (2009). Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. *Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST*, *1*(212), 59–66. https://doi.org/10.1145/1643928.1643943

Roth, D., Lugrin, J. L., Galakhov, D., Hofmann, A., Bente, G., Latoschik, M. E., & Fuhrmann, A. (2016). Avatar realism and social interaction quality in virtual reality. *Proceedings - IEEE Virtual Reality*, *2016-July*, 277–278. https://doi.org/10.1109/VR.2016.7504761

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-

Taking for Conversation. *Language*, *50*(4), 696. https://doi.org/10.2307/412243

Sengün, S., Santos, J. M., Salminen, J., Jung, S., & Jansen, B. J. (2022). Technological Forecasting & Social Change Do players communicate differently depending on the champion played ? Exploring the Proteus effect in League of Legends. *Technological Forecasting and Social Change*, *177*(August 2021). https://doi.org/10.1016/j.techfore.2022.121556

Shih, M. T., Lee, Y. C., Huang, C. M., & Chan, L. (2023). "A feeling of déjà vu": The Effects of Avatar Appearance-Similarity on Persuasiveness in Social Virtual Reality. *Proceedings of the ACM on Human-Computer Interaction*, *7*(CSCW2). https://doi.org/10.1145/3610167

Shin, M., Jung, S., & Biocca, F. (2019). Computers in Human Behavior The uncanny valley : No need for any further judgments when an avatar looks eerie. *Computers in Human Behavior*, *94*(October 2018), 100–109. https://doi.org/10.1016/j.chb.2019.01.016

Sidenmark, L., & Gellersen, H. (2019). Eye, head and torso coordination during gaze shifts in virtual reality. *ACM Transactions on Computer-Human Interaction*, *27*(1), 1–40. https://doi.org/10.1145/3361218

Smith, H. J., & Neff, M. (2018). Communication behavior in Embodied virtual reality. *Conference on Human Factors in Computing Systems - Proceedings*, *2018-April*, 1–12. https://doi.org/10.1145/3173574.3173863

Straus, S. G., & McGrath, J. E. (1994). Does the Medium Matter? The Interaction of Task Type and Technology on Group Performance and Member Reactions. *Journal of Applied Psychology*, *79*(1), 87–97. https://doi.org/10.1037/0021-9010.79.1.87

Sykownik, P., Karaosmanoglu, S., Emmerich, K., Steinicke, F., & Masuch, M. (2023). VR Almost There: Simulating Co-located Multiplayer Experiences in Social Virtual Reality. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3544548.3581230

Tanenbaum, T. J., Hartoonian, N., & Bryan, J. (2020). "how do i make this thing smile?": An Inventory of Expressive Nonverbal Communication in Commercial Social Virtual Reality Platforms. *Conference on Human Factors in Computing Systems - Proceedings*, 1–13. https://doi.org/10.1145/3313831.3376606

Tõke, T., Selvet, R., Tiri, K., & Järveoja, H. (2024). *Ready Player Me*. Ready Player Me, Inc. https://readyplayer.me/

Van Der Wiel, W., Hasberg, M.-P., Weima, I., & Huiskamp, W. (2010). Concept Maturity Levels Bringing structure to the CD&E process. *Interservice/Industry Training, Simulation, and Education Conference*, *10105*, 1–9.

Van Swol, L. M. (2003). The Effects of Nonverbal Mirroring on Perceived Persuasiveness , Agreement With an Imitator , and Reciprocity in a Group Discussion. *Communication Research*, *30*(4), 461–480. https://doi.org/10.1177/0093650203253318

Voinescu, A., Fodor, L. A., Fraser, D. S., & David, D. (2020). Exploring attention in vr: Effects of visual and auditory modalities. *Advances in Intelligent Systems and Computing*, *1217 AISC*, 677–683. https://doi.org/10.1007/978-3-030-51828-8_89

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, L., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., & others. (2004). MEC Spatial Presence Questionnaire (MEC-SPQ). *Report to the European Community, Project Presence: MEC (IST-2001-37661)*, *18*(June), 15. https://doi.org/10.13140/RG.2.2.26232.42249

Weibel, D., Stricker, D., Wissmath, B., & Mast, F. W. (2010). How Socially Relevant Visual

Characteristics of Avatars Influence Impression Formation. *Journal of Media Psychology*, *22*(1), 37–43. https://doi.org/10.1027/1864-1105/a000005

What is Brainstorming? (2016). Interaction Design Foundation - IxDF. https://www.interaction-design.org/literature/topics/brainstorming

Wolbers, J. (2024). *RadishNet*. GitHub. https://github.com/radishnet

Yassin, M., El Antably, A., & Abou El-Ela, M. A. S. (2021). The others know the way: A study of the impact of co-presence on wayfinding decisions in an interior virtual environment. *Automation in Construction*, *128*(April), 103782. https://doi.org/10.1016/j.autcon.2021.103782

Zenner, A., Ullmann, K., Ariza, O., Steinicke, F., & Krüger, A. (2023). Induce a Blink of the Eye: Evaluating Techniques for Triggering Eye Blinks in Virtual Reality. *Conference on Human Factors in Computing Systems - Proceedings*, 1–8. https://doi.org/10.1145/3544548.3580888

Zhang, X., Pi, Z., Zhao, R., Bai, H., Hu, W., Wei, X., Cai, N., & Zhang, L. (2022). Task motivation enhances creative performance in online groups, but not interpersonal interaction. *Interactive Learning Environments*. https://doi.org/10.1080/10494820.2022.2061008

Appendices

a. Counterbalancing Table

Group	Initial Avatar State	Final Avatar State	Initial Room Set
1	Low	High	A
2	High	Low	A
3	Low	High	A
4	High	Low	A
5	Low	High	A
6	High	Low	B
7	Low	High	B
8	High	Low	B
9	Low	High	B
10	High	Low	B

The greyed out rows are groups that were planned for, but cancelled.

b. Survey Questions & Modifications

Measure	Variable Examined	Questions Applicable	Altered Questions	RQ
Networked Minds Measure of Social Presence (Harms & Biocca, 2004)	Co-Presence	1. I noticed (my partner) 2. (My partner) noticed me 3. (My partner's) presence was obvious to me. 4. My presence was obvious to (my partner). 5. (My partner) caught my attention. 6. I caught (my partner's) attention.	1. I noticed (my partners) 2. (My partners) noticed me 3. (My partners') presence was obvious to me. 4. My presence was obvious to (my partners). 5. (My partners) caught my attention. 6. I caught (my partners') attention.	A
	Perceived Message Understanding	1. My thoughts were clear to (my partner). 2. (My partner's) thoughts were clear to me. 3. It was easy to understand (my partner). 4. (My partner) found it easy to understand me. 5. Understanding (my partner) was difficult. 6. (My partner) had difficulty understanding me.	1. My thoughts were clear to (my partners). 2. (My partners') thoughts were clear to me. 3. It was easy to understand (my partners). 4. (My partners) found it easy to understand me. 5. Understanding (my partners) was difficult. 6. (My partners) had difficulty understanding me.	A
	Attentional Allocation	1. I was easily distracted from (my partner)	1. I was easily distracted from (my partners) when other	A

		when other things were going on. 2. (My partner) was easily distracted from me when other things were going on. 3. I remained focused on (my partner) throughout our interaction. 4. (My partner) remained focused on me throughout our interaction. 5. (My partner) did not receive my full attention. 6. I did not receive (my partner's) full attention.	things were going on. 2. (My partners) was easily distracted from me when other things were going on. 3. I remained focused on (my partners) throughout our interaction. 4. (My partners) remained focused on me throughout our interaction. 5. (My partners) did not receive my full attention. 6. I did not receive (my partners') full attention.	
	Perceived Affective Understanding	1. I could tell how (my partner) felt. 2. (My partner) could tell how I felt. 3. (My partner's) emotions were not clear to me. 4. My emotions were not clear to (my partner). 5. I could describe (my partner's) feelings accurately.	1. I could tell how (my partners) felt. 2. (My partners) could tell how I felt. 3. (My partners') emotions were not clear to me. 4. My emotions were not clear to (my partners). 5. I could describe (my partners') feelings accurately. 6. (My partners) could describe	A

		6. (My partner) could describe my feelings accurately.	my feelings accurately.	
MEC-SPQ (Vorderer et al., 2004)	Spatial Situation Model	1. I was able to imagine the arrangement of the spaces presented in the [medium] very well. 2. I had a precise idea of the spatial surroundings presented in the [medium] 3. In my mind's eye, I was able to clearly see the arrangement of the objects presented/described. 4. I was able to make a good estimate of the size of the presented space. 5. I was able to make a good estimate of how far apart things were from each other. 6. Even now, I still have a concrete mental image of the spatial environment. 7. Even now, I could still draw a plan of the spatial environment	[medium] is virtual environment	B

		in the presentation. 8. Even now, I could still find my way around the spatial environment in the presentation.		
	Self-Location	1. I had the feeling that I was in the middle of the action rather than merely observing. 2. I felt like I was a part of the environment in the presentation. 3. I felt like I was actually there in the environment of the presentation. 4. I felt like the objects in the presentation surrounded me. 5. It was as though my true location had shifted into the environment in the presentation. 6. It seemed as though my self was present in the environment of the presentation. 7. I felt as though I was		B

		physically present in the environment of the presentation. 8. It seemed as though I actually took part in the action of the presentation.		
Computer Brainstorms Questionnaire Measures (Dennis & Valacich, 1993)	Satisfaction	1. How do you feel about the process by which you generated ideas? 2. How do you feel about the idea proposed? 3. All-in all, how did you feel?	1. I feel satisfied with the idea generation process I used. 2. I feel satisfied with the ideas proposed. 3. All-in all, I feel satisfied.	C
	Production Blocking	1. When you thought of an idea… (immediately) 2. Did you express your ideas… (immediately)	1. When I thought of an idea, I could express it immediately. 2. I expressed my ideas soon after thinking of them.	C
	Evaluation Apprehension	1. Did you feel any apprehension about generating your ideas? 2. How at ease were you during the idea generation session?	1. I did not feel a lot of apprehension while generating my ideas. 2. I felt at ease during the idea generation session.	C
	Free Riding	1. How much do you feel you participated in this idea generation session?	1. I feel like I participated in the idea generation session. 2. I feel satisfied with my own	C

		2. How satisfied are you with your own performance on this task?	performance on this task.	
	Synergy and Stimulation	1. How stimulating did you find this task? 2. How interesting was this idea generation task? 3. How motivated were you to generate quality ideas?	1. I found this task stimulating. 2. This idea generation task was interesting. 3. I was motivated to generate quality ideas.	C
	Sufficient Time	1. For this idea generation session, did you… 2. Considering all the ideas you thought of, did you…	1. I had as much time as I needed for the session. 2. I did not have enough time to express all my ideas.	C

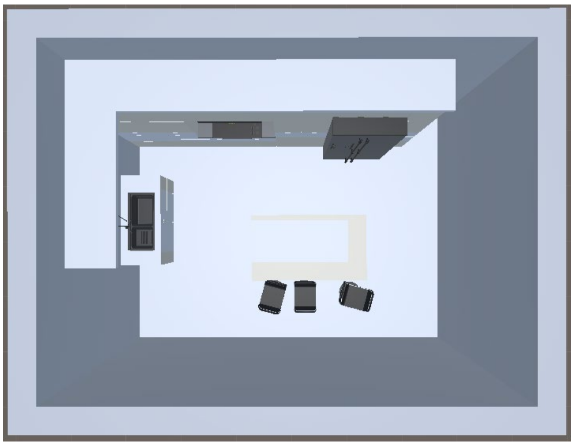
c. Ad Hoc Questions

Question	Input	RQ Related
What features of the room did you consider to be an advantage?	Free Response	C
What features of the room did you consider to be a disadvantage?	Free Response	C
What recommendations do you have for the environment?	Free Response	C
I generally liked the environment I evaluated.	1-7	D
The environment is well designed.	1-7	D
It was difficult to come up with disadvantages about the room environment.	1-7	D

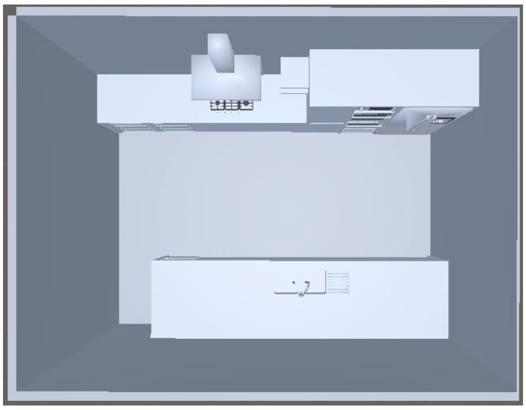
d. Kitchen Sets

Note: The related kitchens had ceilings covering them when used. In the case of B2, there were walls like the others present.

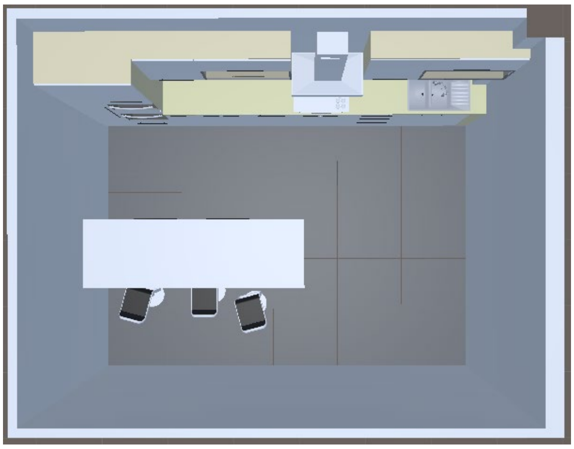
A1



A2



B1



B2

e. Internal Review Board (TNO Ethics Committee) Completion Statement

) TNO Intern

innovation for life

Memo

www.tno.nl
jan.vanerp@tno.nl
+31613738409

Datum
26 april 2024
Onze referentie
2024-044

Aan	J. Siri Dr. W.M. Post
Kopie aan	Drs. C. Jansen
Van	J.B.F. (Jan) van Erp
Onderwerp	Positief Advies 2024-044 Impacts of Avatar Fidelity in Multi-User Virtual Reality Brainstorming for Room Design

Background
On 16 April 2024 the research proposal "Impacts of Avatar Fidelity in Multi-User Virtual Reality Brainstorming for Room Design" was submitted to the TNO Institutional Review Board (IRB). The composition of the IRB is specified on the TNO Intranet page 'Human Research'.

Advice
The IRB had considered the proposed research on the basis of its regulations and expresses a positive recommendation.
The advice is determined in accordance with the methodology that can be found on the TNO Intranet page "Human Research".

In its deliberations, the IRB has considered the research design and privacy aspects, in addition to – where relevant – the ethical aspects and the burden and the risks to the research participants.
In the event of important modifications to the research or in the event incidents occur, the project leader shall inform the IRB. This may lead to amended recommendations.

Sincerely,
On behalf of the IRB,

Jan van Erp,
Chair Institutional Review Board TNO

Leiden, 26 April 2024

The Review Board has based its deliberations on the following submitted documents:
- Application form (16-04-2024)
- Research plan (16-04-2024)
- Participant Information Form
- Notification /Certificate of Insurance
- Quick scan DPIA

) TNO Intern 1/1

f. Room Survey

Open Ended: In English, please write out your responses in the spaces below.

What features of the room did you consider to be an advantage?

What features of the room did you consider to be a disadvantage?

What recommendations do you have for the environment?

Ratings: Please fill the bubble corresponding to your level of agreement.

I generally liked the environment I evaluated.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

The environment is well designed.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

It was difficult to come up with disadvantages about the room.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

g. Extended Room Survey

Open Ended: In English, please write out your responses in the spaces below.

What features of the room did you consider to be an advantage?

What features of the room did you consider to be a disadvantage?

What recommendations do you have for the environment?

Ratings: Please fill the bubble corresponding to your level of agreement.

I did not have enough time to express all my ideas.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree — Neutral — Strongly Agree

(My partners) remained focused on me throughout our interaction.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree — Neutral — Strongly Agree

All-in all, I feel satisfied.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree — Neutral — Strongly Agree

It seemed as though my self was present in the environment of the presentation.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

My thoughts were clear to (my partners).

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

Even now, I could still find my way around the spatial environment in the presentation.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I did not receive (my partners') full attention.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

When I thought of an idea, I could express it immediately.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I feel satisfied with the ideas proposed.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

In my mind's eye, I was able to clearly see the arrangement of the objects presented/described.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I did not feel a lot of apprehension while generating my ideas.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

Understanding (my partners) was difficult.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners') presence was obvious to me.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners) caught my attention.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

The environment is well designed.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners) could describe my feelings accurately.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I was motivated to generate quality ideas.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I felt at ease during the idea generation session.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I caught (my partners') attention.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners') thoughts were clear to me.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I felt as though I was physically present in the environment of the presentation.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I remained focused on (my partners) throughout our interaction.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

(My partners) noticed me.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

Even now, I could still draw a plan of the spatial environment in the presentation.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

I was able to make a good estimate of the size of the presented space.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

My presence was obvious to (my partners).

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

I felt like I was actually there in the environment of the presentation.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

I feel satisfied with the idea generation process I used.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

I expressed my ideas soon after thinking of them.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

I generally liked the environment I evaluated.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree | Neutral | Strongly Agree

(My partners') emotions were not clear to me.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

(My partners) could tell how I felt.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

I feel like I participated in the idea generation session.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

I had a precise idea of the spatial surroundings presented in the VR environment.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

I was able to make a good estimate of how far apart things were from each other.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

It was as though my true location had shifted into the environment in the presentation.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

(My partners) had difficulty understanding me.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

My emotions were not clear to (my partners).

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

I had as much time as I needed for the session.

○ ○ ○ ○ ○ ○ ○

Strongly Disagree Neutral Strongly Agree

I felt like I was a part of the environment in the presentation.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I feel satisfied with my own performance on this task.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

This idea generation task was interesting.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I could tell how (my partners) felt.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners) did not receive my full attention.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

It was difficult to come up with disadvantages about the room.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I found this task stimulating.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I was able to imagine the arrangement of the spaces presented in the [medium] very well.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners) found it easy to understand me.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

Even now, I still have a concrete mental image of the spatial environment.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I felt like the objects in the presentation surrounded me.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I noticed (my partners).

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I was easily distracted from (my partners) when other things were going on.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

(My partners) was easily distracted from me when other things were going on.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I could describe (my partners') feelings accurately.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

I had the feeling that I was in the middle of the action rather than merely observing.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

It seemed as though I actually took part in the action of the presentation.

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree

It was easy to understand (my partners).

○ ○ ○ ○ ○ ○ ○
Strongly Disagree Neutral Strongly Agree