Effects of Avatar Transparency on Social Presence in Task-centric Mixed Reality Remote Collaboration

Boram Yoon, Jae-eun Shin, Hyung-il Kim, Seo Young Oh, Dooyoung Kim and Woontack Woo

Fig. 1: This study investigates how the degree of remote avatar transparency affects user experience during task-centric dynamic MR collaboration. A local AR user and remote VR user collaborated in a shared virtual space to put together space research bases while being exposed to three different types of avatars: (A) Nontransparent; (B) Semi-transparent; and (C) Near-transparent.

Abstract—Despite the importance of avatar representation on user experience for Mixed Reality (MR) remote collaboration involving various device environments and large amounts of task-related information, studies on how controlling visual parameters for avatars can benefit users in such situations have been scarce. Thus, we conducted a user study comparing the effects of three avatars with different transparency levels (Nontransparent, Semi-transparent, and Near-transparent) on social presence for users in Augmented Reality (AR) and Virtual Reality (VR) during task-centric MR remote collaboration. Results show that avatars with a strong visual presence are not required in situations where accomplishing the collaborative task is prioritized over social interaction. However, AR users preferred more vivid avatars than VR users. Based on our findings, we suggest guidelines on how different levels of avatar transparency should be applied based on the context of the task and device type for MR remote collaboration.

Index Terms—Telepresence, Avatars, Mixed Reality, Augmented Reality, Virtual Reality, Collaboration, Embodiment

1 Introduction

In response to the growing demand for remote collaboration solutions, interest in immersive 3D telepresence systems that employ Augmented Reality (AR), Virtual Reality (VR), and Mixed Reality (MR) technologies has also been on the rise. These systems enable distant users to communicate while performing various types of tasks together, such as face-to-face collaborative discussion [2, 20, 38, 47, 63] or collaboration for accomplishing the given mission [10, 11, 35, 41, 44], by interacting with virtual elements in 3D spaces. Among these elements, the representation of avatars as embodiments of collaborating users plays a crucial role in determining user experience and task performance [37]: enhancing mutual understanding or interfering with intuitive information acquisition due to their presence in the collaboration space.

Since a remote avatar is an essential collaboration component that moves with autonomy, the significance of how avatars are visualized increases in situations where users are required to utilize large amounts of various information and move actively around the shared virtual space through various types of devices. Unlike user-centric situations—the main focus is on collaborating users [2, 23, 55, 64, 65], task-centric situations—the main focus is the completion of the task through collaboration—may face different challenges in visualizing remote users to less disturb local users using virtual information.

Adjusting the level of transparency can be one way to resolve such a situation where the task-related elements require more attention than the remote avatar, as this would be beneficial for hiding attention-grabbing glitches, avoiding occlusion, and reducing crowded feeling [1, 9, 46, 59]. Moreover, as the shared space becomes wider and visually complex, differences in device environments can become more apparent [25, 27, 49, 50, 53]. In such cases, manipulating the visibility of avatars could help minimize these effects of devices, thereby raising social presence and improving task performance. However, prior works on remote avatar representation have mostly focused on generating avatars [13, 62, 65], visualizing partial bodies [2, 8], or its realistic appearances [23, 63]. Research on how the degree of avatar transparency affects remote collaboration in asymmetrical MR settings has been scarce [56].

For this, we compared the effect of remote avatar transparency in three levels (Nontransparent, Semi-transparent, and Near-transparent) among two users in different MR device environments (AR and VR) during interactive task-centric remote collaboration (Fig. 1). Factors of social presence, subjective perception, and task performance were evaluated: We found that participants felt no difference between the three avatar types in terms of collaboration involvement, attention, negative perception, and task performance. However, the Near-transparent avatar led to lower social awareness, whereas the Nontransparent avatar induced participants to avoid close encounters. The Semi-transparent avatar was perceived better than the Near-transparent in terms of intimacy and preference. Between different devices, VR users experienced higher social presence and action possibilities, while AR users felt greater task load but more positive feelings for their partner.

Based on these findings, we derive implications for applying different avatar transparency levels to foster better task-centric remote MR collaboration. Consequently, the main contributions of this study are:

- Investigating remote avatar transparency, which has not been evaluated previously, especially in asymmetric MR configuration.
- Suggesting implications for context-adaptive applications of avatar transparency to support task-centric MR remote collaboration.

- *Boram Yoon, Hyung-il Kim, Seo Young Oh and Dooyoung Kim are with KAIST UVR Lab. E-mail: {boram.yoon1206 | hyungil | seoyoung.oh | dooyoung.kim}@kaist.ac.kr.*
- *Jae-eun Shin is with KAIST KI-ITC ARRC. E-mail: jaeeunshin@kaist.ac.kr.*
- *Woontack Woo is with KAIST UVR Lab and KAIST KI-ITC ARRC. Corresponding Author. E-mail: wwoo@kaist.ac.kr.*

Manuscript received 25 March 2023; revised 17 June 2023; accepted 7 July 2023. Date of publication 2 October 2023; date of current version 31 October 2023. This article has supplementary downloadable material available at https://doi.org/10.1109/TVCG.2023.3320258, provided by the authors. Digital Object Identifier no. 10.1109/TVCG.2023.3320258

2 Related Work

2.1 Social Presence and Avatar Transparency in MR Remote Collaboration

Social presence is regarded as a key factor in assessing the user experience of remote collaboration systems. Defined as "the sense of being together" [7] and "the sense of communicating with each other" [22], social presence conceptualizes the essence of telepresence and remote collaboration in that the collaborating partners do not physically coexist but are virtually represented to one another via avatars. In achieving high levels of social presence during 3D remote collaboration, the visual appearance of remote avatars therefore plays an important role.

The majority of studies on the visual presentation of virtual avatars have been mostly focused on comparing the effects of different partial body visualizations [2, 5, 21, 63] and character styles such as level of realism [23, 64] and fidelity [13, 45, 62, 65] on social presence and user perception. Studies on partial body representation commonly found that the elimination of the body as a whole or in parts led to lower social presence: Depending on the collaboration context, it is important to provide visual information on body parts that are involved in delivering communicational cues, such as body joints or the torso [2, 15, 63]. In the case of character styles, low visual fidelity did not always have a negative effect because other factors such as task type, kinetic fidelity, and user resemblance were simultaneously involved [13, 45, 65].

These findings provide the grounds on which avatar transparency, which has been an under-explored topic for avatar representation in MR, should be brought into focus. Unlike partial avatars, transparent avatars provide continuous body images that can communicate body cues fully while maintaining a lower visual presence. In addition, they are less prone to occlusion than opaque, low-fidelity avatars and are suitable in collaboration scenarios where the presence of the other user is required but should not be impeding the task [46, 59].

Some previous works attempted to leverage avatar transparency in enhancing user experience in 3D remote collaboration. Transparent avatars were implemented with visible outlines in developing remote MR collaboration systems [47, 55]. Shadow-Avatars [26] presented transparent avatars to avoid collision between co-located users in VR. Other works suggested the use of transparent avatars in multi-user 3D remote collaboration to decrease obstruction and occlusion in the users' Field of View (FoV): SocialSlider [59] employed transparent avatars in proposing interaction methods for social VR, and Weissker and Froehlich [57] applied the concept to support obstruction-free group navigation in distributed virtual environments. However, these works did not evaluate and prove the effect of transparency on user experience.

Few works have made a direct comparison between different levels of avatar transparency on user perception in various 3D environments. Martini et al. [31] examined the effects of avatar transparency on self-body illusion in VR. They found that when the transparency level increased, users' body ownership decreased. Peck et al. [42] observed how perceived humanness was affected based on four transparency levels among virtual characters with different skin tones through an AR Optical See-Through Head Mounted Display (OST-HMD). They revealed that dehumanization occurred when the avatar's transparency increased. While these studies commonly found that increasing transparency had negative effects on image perception in both AR and VR, they neither evaluated remote embodiment such as social presence nor were conducted in real-time 3D remote collaboration.

2.2 Social Presence and Device Environment in MR Remote Collaboration

MR, within Milgram and Kishino's Reality-Virtuality Continuum [33], refers to an environment that consists of a blending of real and virtual worlds, allowing users to be aware of and interact with both real and virtual objects simultaneously: The concept of collaborative MR extends this term to telepresence systems by considering artificiality, transportation, and shared spatiality [6]. MR remote collaboration generally features an asymmetric setup [12, 52], where an AR user interacts with a remote avatar in a virtual space overlaid on reality, while a VR user perceives a real user's virtual avatar within an immersive shared space. Along with these fundamental differences, hardware features such as resolution and FoV result in further discrepancies between the experience of users collaborating through each device [25, 50], namely social presence: Reducing this gap between users during asymmetric MR collaboration is a crucial challenge.

Despite this importance, only a small number of studies have explored how different MR environments affect social presence, especially for remote collaboration. Grandi et al. [14] and Li et al. [29] evaluated the collaboration experience of two co-located users each using a VR and mobile AR device. While the former found that the AR user moved less actively due to the AR device's narrower FoV, the latter suggested that supplying avatars and more visual cues for AR users may help overcome this disadvantage and raise social and spatial presence. However, their user tasks were limited to a confined area with little movement and the study was not done in an HMD-based MR collaboration context. In addition, Piumsomboon et al. [43, 44] observed that the AR HMD's limited FoV may have led to a lower co-presence for its user than the VR user. Bai et al. [3] also found that VR users felt higher spatial presence than AR users because they could be more aware of their surroundings and walk around freely. Teo et al. [54] claimed that to compensate for the narrower FoV of AR HMD during asymmetric MR remote collaboration, a stronger visual presence of the remote VR user is required. However, the disadvantages caused by the AR HMD's FoV have not been systematically investigated for avatar-mediated collaboration in these works.

Shin et al.'s study also asserted that the limited AR HMD's FoV could induce more difficulties in finding and perceiving their virtual partner and shared space; The VR user felt higher action possibility due to more agency in moving around the overall virtual space under a wider and clearer VR display [49]. Furthermore, Rhee et al. [47] found that the difference in the FoVs led to the spontaneous distinction between roles that the AR and VR users each took on: VR users tended to lead the situation as they had a better understanding of the visual elements. Some other studies pinpointed the visual perception of avatars over various modes of MR. Wolf et al. found that the sense of acting in a virtual space and self-body estimation was lower through an AR OST-HMD than other HMDs, while environmental involvement or avatar attractiveness was not different [60, 61]. A study on user preference for the partner's avatar recommended holographic avatars for AR users because their transparency enables users to obtain visual elements more easily [40]. However, these studies have not explored how collaborating users perceive the other person's avatar through different devices.

Jo et al. [23] claimed that the remote partner in the real background increased co-presence compared to the virtual during conversation; Their study used the same VST-HMDs, so further analysis for asymmetrical setups is still needed. While Yoon et al. [64] focused on how AR and VR led to differing perceptions of the remote collaborator, it was also limited to hand representation in static collaboration where the occlusion of visual information did not occur. Whereas AR users felt higher social presence and lower workload because seeing their real hands made communication easier, VR users reported higher load due to occasional misrepresentations of the wholly virtual hands they were shown. In this study, we attempt to build upon these works to address the gap by focusing on how the inherent differences in the AR and VR devices affect user collaboration through avatars with different transparency levels in a task-centric collaboration involving large amounts of task-related information and active movement.

3 Methodology

For this study, the following research questions were asked:

- RQ1. How does avatar transparency affect users in terms of social presence, perception, and task performance in task-centric MR remote collaboration?
- RQ2. How does the type of mediation affect users in terms of social presence, perception, and task performance in task-centric MR remote collaboration?
- RQ3. How do users involved in task-centric MR remote collaboration perceive, interact, and behave with respect to avatar transparency and type of mediation?

3.1 Experimental Conditions and Hypotheses

For avatar transparency, we set three conditions: (1) Nontransparent (Non), (2) Semi-transparent (Semi), and (3) Near-transparent (Near) (Fig. 2(A)). To determine each level, we referred to previous studies [31, 32, 42] and conducted a pilot test to set values that were distinguishable in the experimental setting. The Nontransparent avatar, representing the default, fully opaque avatar (0% transparency) was set as the first condition. Based on prior works, the level of transparency for the Semi-transparent avatar was set at 50%. The last condition maintained the minimum level of presence, its transparency level set at 85%. For types of mediation, we set two conditions: (1) AR HMD (the local side), and (2) VR HMD (the remote side) (Fig. 2(B)).

We summarized the implications based on the findings from earlier works to be followed up in our study: Related to avatar representation, the more incomplete the avatar is shown, such as a partial or transparent body, the lower the user's social presence. Moreover, although the avatar's low visual fidelity does not always induce negative effects, a more realistic and complete representation results in higher embodiment in both self-body and remote body perception and delivers a better understanding of their surroundings. Regarding the AR/VR devices, recent studies investigating avatar-mediated MR remote collaboration revealed that AR HMD users felt higher social presence because they had seen a real environment and their real bodies. However, since the VR HMDs provide a wider and more vivid display, the VR HMD users generally take more active movement and a leading role—it even may deliver a burden—with a better understanding of the entire situation during MR remote collaboration. In line with this, four sets of hypotheses were drawn:

- H1-1. Factors of social presence in a task-centric MR collaboration will be measured lower for users collaborating with the Near-transparent remote avatar.
- H1-2. Factors of social presence in a task-centric MR collaboration will be measured higher for users wearing an AR HMD than those wearing a VR HMD.
- H2-1. Perceived action possibility in a task-centric MR collaboration will be measured higher for users collaborating with the Non-transparent remote avatar.
- H2-2. Perceived action possibility in a task-centric MR collaboration will be measured higher for users wearing a VR HMD than those wearing an AR HMD.
- H3-1. Interpersonal impression in a task-centric MR collaboration will be measured higher for users collaborating with the Nontransparent remote avatar.
- H3-2. Interpersonal impression in a task-centric MR collaboration will be measured higher for users wearing an AR HMD than those wearing a VR HMD.
- H4-1. Task load in a task-centric MR collaboration will be measured higher for users collaborating with the Near-transparent remote avatar.
- H4-2. Task load in a task-centric MR collaboration will be measured higher for users wearing a VR HMD than those wearing an AR HMD.

3.2 Study Design and Task

The experiment employed a 3×2 mixed factorial design with avatar transparency as the within-subject factor and device type as the between-subject factor: The pair of participants were randomly assigned to either an AR or VR HMD and were exposed to all three avatar transparency conditions. The order of avatar transparency condition was counter-balanced based on a Latin Square method to avoid ordering effects. For our experimental task, we set the following design rationale:

- DR1. The main goal is to accomplish each user's personal mission during an MR collaboration, which assumes a task-centric situation.
- DR2. While each participant is given individual missions, the task can only be completed through the users' collaboration.
- DR3. The task should ensure both users' active and constant movement across the shared space so that they are sufficiently exposed to their collaborator's avatar.

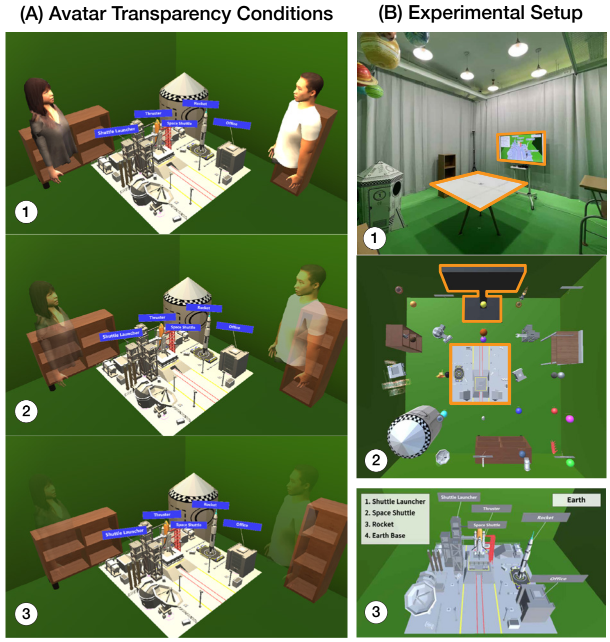

Fig. 2: (A) The three avatar transparency conditions: (1) Nontransparent, (2) Semi-transparent, and (3) Near-transparent; (B) The experimental setups for two device conditions and our main task instruction: (1) AR, (2) VR and virtual elements of the Earth task (The orange outline indicates the shared workspace (a table) and TV screen), and (3) An example of the virtual space research base (the Earth task)

- DR4. The task should ensure sufficient overlaps between the remote avatar and task environment—task-related and unrelated physical and virtual elements—as much as possible.
- DR5. Various types of information including text, 2D images, and 3D objects should be utilized for the collaborative task.

To satisfy this, we selected a virtual space research base building task where users on both sides were each given different parts to find and manipulate to complete the whole scene. Because we have the three levels of transparency factor, three different space research bases—a Mars (Fig. 1(B)), a Moon (Fig. 1(A)), and an Earth (Fig. 1(C))—were designed. Based on the first rationale, the two participants in a paired group were provided different types of instructions (Type A or B) that they should follow, which were shown on a TV screen in the shared space. Each space base consisted of four main components including a building, machines or vehicles, and a base environment.

Participants had to complete these components through three types of actions: (1) Assembly and placement, (2) Coloring and texturing, and (3) Tagging. According to DR1, we tried to exclude any direct instructions that induce forced interaction between the collaborators, such as finding or handing over objects to the partner. Some of the instructions could only be achieved after the partner finished theirs; By creating situations where a user could not proceed without knowing how far along the task the partner had come, we attempted to prevent participants from working by themselves without any communication with the other (DR2). Although the given instructions did not include direct interaction between the users, they were asked at the beginning of the task to bow and shake hands with the other's avatar. When they completed the task, they were required to give each other a high-five.

In setting the collaboration space based on DR3 to DR5, the size of the shared space was fixed at 3.7m × 3.7m, providing sufficient space for the participants to walk freely while allowing the remote avatar to be recognized within the FoV during the collaboration. Seven furniture items—a touchscreen TV monitor, a table, two bookshelves, a lectern, a 3D paper rocket, and a toy model of the solar system on the ceiling—were used to represent the collaborative space. For the collaboration task, the base frame of each task was placed on the square table in the shared space, and the TV screen was used to display each participant's different instructions. The AR participant directly touched the real TV monitor to turn the instruction pages, and the VR participants did the same work by interacting with the virtual TV replicating the real one.

To maintain the same level of task difficulty between task sets and instructions, the number of following components were controlled: (1) augmented items in the scene (32 or 33 items), (2) to-be-assembled objects (four objects), and (3) required interactions with objects (17 instructions). Also, the task-related virtual elements—15 or 16 space base components, 12 material balls, and five name tags—were augmented in the shared scene. Their default positions were randomly determined but evenly distributed across the space, considering the average knee height and eye level of the recruited ethnic group. Within each session, the order of the three task sets was also counterbalanced based on the Latin Square method. The type of instructions was randomly assigned to each participant, and the task was symmetric so that the paired participants had an equal role for collaboration.

3.3 Implementation and Setup

We designed and implemented an asymmetric MR remote collaboration system using the Unity game engine, version 2020.3.19f1. The host AR user wore a Microsoft HoloLens 2[1] and the remote VR user wore a Meta Quest 2[2], and they were connected in real-time. The Photon Unity Networking 2[3] plugin was utilized to synchronize each user's head and hands poses, the shared virtual object's position and rotation, and any updates resulting from the task application functions in the shared scene. The Mixed Reality ToolKit version 2.7.2[4] was utilized to implement the object manipulation functions and the graphical user interface. The three task actions were based on direct bare-hand manipulation. Avatar motion was generated based on an Inverse Kinematics method with HMD-tracked head and hands motion data by using the FinalIK plugin[5].

Since our study was not focused on the aesthetics of the remote avatar style and to prevent delivering unintended moods implied by other avatar styles [56,63], we used the same realistic 3D human models from Renderpeople[6] for all participants but only matched their gender and ethnicity. We adopted only the upper body, regarded as a sufficient model [56,63,65], considering the limited AR HMDs' FoV and the focused task area. To represent the transparent avatars, we implemented a custom shader, which uses a two-way depth pass utilizing a z-buffer and texture alpha blending sequentially for handling alpha sorting. The level of avatar transparency was mapped to alpha values ranging from 0.0 to 1.0—0.5 for the Semi-transparent and 0.85 for the Near-transparent avatar. We also modified the brightness of the avatars' skin tones and clothes to minimize the effect of real lighting on their appearance in the AR environment. Moreover, to lessen further effects between the task objects' colors and transparent avatars, especially more problematic in the OST-HMD, we utilized limited colors—greyscale, or the primary and secondary colors in the additive model—and textures.

The study was set up in an empty studio, and the AR host space and the remote VR space were set up side by side. Each space occupied the same size as the shared space and was separated by curtains. The participants' real voices were used for verbal communication, which has often been configured in previous studies on MR remote collaboration [24,35,63,64]. The task space's lighting was also controlled in advance: The AR space was blocked by blackout curtains in all directions to reduce the external light impact. The controllable smart lighting system implemented in the studio was controlled through the sessions with the same level of illuminance based on the pilot testing measured through the two mobile light meters. The brightness of HoloLens 2 was also always set to 100% during the session.

[1]https://www.microsoft.com/en-us/hololens/
[2]https://store.facebook.com/quest/products/quest-2
[3]https://www.photonengine.com/PUN
[4]https://github.com/Microsoft/MixedRealityToolkit-Unity
[5]https://assetstore.unity.com/packages/tools/animation/final-ik-14290
[6]https://renderpeople.com/

3.4 Dependent Variables

We measured social presence, perceived action possibility, interpersonal impression, task load and completion time. As MR remote collaboration systems have been defined and investigated based on three major components—Environment, Avatars, and Interaction [48], we used three broadly adopted social presence measurements for multifaceted analysis within diverse sub-concept and perspectives of each tool [39].

The Networked Minds Measure of Social Presence (NM) by Harms and Biocca [16] is focused on the sense of being with another, attention, and mutual understanding under interaction. The four subscales—co-presence, attentional allocation, perceived message understanding, and perceived behavioral interdependence—were adapted regarding the study purpose; 24 items on a seven-point Likert scale were measured. The Temple Presence Inventory (TPI) by Lombard et al. [30] focuses on the medium and content in conveying telepresence; 10 seven-point Likert scale items in two subscales—actor within medium and engagement—were evaluated. From the Nowak's Social Presence (NSP) [36], which includes subfactors for between-user intimacy and environmental immersion, 11 items on a five-point Likert scale for Co-presence and four items on a seven-point Likert scale for Telepresence were utilized.

To investigate how avatar transparency and device types affect users' subjective perception of their ability to manipulate tasks, as well as their emotional feelings during dynamic movements and encounters, the following factors were used: Perceived action possibility, one of the subfactors in the Spatial Presence Experience Scale (SPES) [19], was evaluated with four seven-point Likert scale items. Interpersonal impression was comprehensively analyzed by five items on emotional reaction proposed by Bailenson et al. [4] and a single item that evaluated likability. Both factors were rated on a seven-point Likert scale.

Next, the NASA Task Load Index (NASA TLX) [17,18] was used to assess perceived workload. The 'raw TLX' approach was applied for self-evaluated scores ranging from 0 to 100. For task completion time, lastly, we recorded and calculated in seconds the time it took from the participant's first high-five at the beginning of the task to the ending high-five upon its completion.

3.5 Study Procedures

The study content and procedures were approved by an Institutional Review Board and followed COVID-19 safety protocols. First, participants agreed to the terms of the experiment and filled out a demographic questionnaire. We then explained the study process and details about the experiment task. The participants were informed that they could freely communicate with each other during the task. Additionally, they were asked to refrain from doing the partner's work and to only provide assistance when the other needed it (DR2).

As our main task involved bare-hand interactions with virtual objects, we held brief practice sessions before the actual task: In another virtual space set up specifically for this where the users were not connected, the participants learned how to operate the TV monitor and follow the instructions. Once they prepared for the main session after sufficient practice was delivered, each participant moved to a preassigned standing position to face each other when they started.

When the main task started and both avatars appropriately appeared, they greeted each other and then began the session by exchanging a high-five. During the main task, we did not limit the time and encouraged participants to move around and communicate freely with one another. When both participants finished each of their instructions, they checked with their partner to confirm that all steps were completed and gave each other a high-five again to confirm the end of the task in that condition.

After finishing each main task, they took off their HMDs to avoid dizziness and answered the post-task questionnaires. The pairs repeated the above main task procedure three times within the different avatars and task sets. In the post-experiment interview, participants answered questions on how the avatars with different transparency levels affected their overall experience. The interview was conducted separately for each participant to prevent them from being affected by their partner. Each study session took approximately 80 to 90 minutes.

3.6 Participants

We recruited a total of 54 participants through the campus website. 34 (63%) of participants identified as female and 20 (37%) as male. Their ages ranged from 20 to 38 years ($M = 26.50$, $SD = 3.58$). All pairs were assigned randomly to one of 27 pairs consisting of an AR user and a VR user and were assumed to be not previously acquainted. Of the 27 pairs, 16 were matched between female and male (59.3%), 9 between female and female (33.3%), and 2 between male and male (7.4%).

The participant's previous experience related to AR/VR HMDs was asked: five of them (9.26%) had never experienced any HMDs, 12 (22.22%) only tried once, and 15 (27.78%) up to five times. Otherwise, 22 (40.74%) had moderate to high levels of experience with HMDs: seven (12.96%) had tried on up to nine times, and 15 (27.78%) had more than 10 times. Their prior experience with 3D remote collaboration systems was also asked: the majority of participants (44, 81.49%) answered that they had never (64.82%) or only once (16.67%) used such systems before. Six of them (11.11%) had a moderate level of experience (up to five times: 9.26%; up to nine times: 1.85%), and only four of them (7.41%) had experienced more than 10 times.

4 Results

For subjective measures, we used a two-way repeated measures ANOVA with the Aligned Rank Transform (ART) proposed by Wobbrock et al. [58] for non-parametric factorial analysis ($\alpha = 0.05$). All post-hoc pairwise comparisons were Bonferroni corrected. The internal consistency among Likert items was tested with the reliability coefficient of Cronbach's alpha. For objective measures, we first examined the normality of data distribution and homogeneity of variances through the Shapiro-Wilk and Mauchly's test of Sphericity. If the normality and homogeneity assumptions were satisfied, we ran a one-way repeated measures ANOVA to compare the difference. Otherwise, we also applied non-parametric ART analysis. We excluded one data point as an outlier in task load due to the participant's incorrect input, and all the first trials in each avatar transparency condition (nine in total) were excluded for analyzing task completion time. The main statistical results are summarized as follows:

- R1. Participants felt no differences between the three avatar conditions on the factors indicating: (1) Collaboration task and experience involvement; (2) Attentional allocation; (3) Emotional reaction; and (4) Task performance.
- R2. The Near-transparent avatar was measured lower than the Semi-transparent or Nontransparent on the factors of social presence concerning awareness and understanding of their partner.
- R3. The Nontransparent avatar resulted in higher perceived behavioral interdependence and fewer close encounters than the Near-transparent avatar.
- R4. The Semi-transparent avatar was perceived better than the Near-transparent on the factors indicating intimacy and closeness, and preference.
- R5. VR users felt higher social presence and action possibility than AR users, and AR users reported higher task load and more positive feelings towards their partners.

4.1 Social Presence

Networked Minds Social Presence (NM): The aggregated NM social presence score, which combines all four sub-scales, showed a reliable level of Cronbach's alpha ($\alpha = .950$). The internal consistency of the four subscales also showed an acceptable value (co-presence: $\alpha = .905$; attentional allocation: $\alpha = .800$; perceived message understanding: $\alpha = .938$; perceived behavioral interdependence: $\alpha = .897$).

We found significant main effects of both transparency ($F(2,130) = 10.819$, $p < .001$, $\eta_p^2 = .143$) and device ($F(1,130) = 34.542$, $p < .001$, $\eta_p^2 = .210$) on the aggregated NM. Post-hoc revealed significant differences between Non and Near ($p < .001$), and Semi and Near conditions ($p = .005$). We found no interaction effect between the two factors ($F(2,130) = 1.942$, $p = .147$, $\eta_p^2 = .029$).

For co-presence, significant main effects were found for both transparency ($F(2,130) = 25.414$, $p < .001$, $\eta_p^2 = .281$) and device ($F(1,130) = 36.164$, $p < .001$, $\eta_p^2 = .218$). The post-hoc tests found significant differences between Non and Near ($p < .001$), and Semi and Near conditions ($p < .001$): Participants collaborating with the Near-transparent avatar reported lower co-presence than with the Semi-transparent or Nontransparent avatar. However, there was no interaction effect between the factors ($F(2,130) = 2.533$, $p = .083$, $\eta_p^2 = .038$).

A significant main effect was only found for device on attentional allocation ($F(1,130) = 17.329$, $p < .001$, $\eta_p^2 = .118$). No significant effect for transparency ($F(2,130) = 1.889$, $p = .155$, $\eta_p^2 = .028$), nor interaction effect was found ($F(2,130) = 1.743$, $p = .179$, $\eta_p^2 = .026$).

There were significant main effects of both transparency ($F(2,130) = 6.662$, $p = .002$, $\eta_p^2 = .093$) and device ($F(1,130) = 26.578$, $p < .001$, $\eta_p^2 = .170$) for perceived message understanding. In the post-hoc test, we found significant differences in the pairs of the Non and Near ($p = .003$), and Semi and Near conditions ($p = .012$): Participants felt more difficulties in understanding messages from the Near-transparent avatar representing their partner than the Semi-transparent or Nontransparent. No significant interaction effect was found between the factors ($F(2,130) = .602$, $p = .549$, $\eta_p^2 = .009$).

The perceived behavioral interdependence showed significant main effects of both transparency ($F(2,130) = 3.964$, $p = .021$, $\eta_p^2 = .057$) and device ($F(1,130) = 15.670$, $p < .001$, $\eta_p^2 = .108$). The pairwise comparisons found a significant difference between Non and Near ($p = .024$): Participants experienced that their behavior was less affected by the remote avatar's action with the Near-transparent avatar compared to the Nontransparent avatar. An interaction effect between the two factors was also not found ($F(2,130) = 1.583$, $p = .209$, $\eta_p^2 = .024$).

Temple Presence Inventory (TPI): We deleted one item in the subscale engagement because the internal consistency was not satisfied with this item. The aggregated TPI ($\alpha = .888$) and each of the two subscales all showed an acceptable value of Cronbach's alpha (actor within medium: $\alpha = .892$; engagement: $\alpha = .725$).

For the aggregated TPI, significant main effects were found for both transparency ($F(2,130) = 10.489$, $p < .001$, $\eta_p^2 = .139$) and device ($F(1,130) = 35.347$, $p < .001$, $\eta_p^2 = .214$). The post-hoc test showed that Non and Near ($p < .001$), and Semi and Near ($p = .002$) conditions were significantly different. No significant interaction effect between the two factors was found ($F(2,130) = 1.924$, $p = .150$, $\eta_p^2 = .029$).

We found significant main effects of both transparency ($F(2,130) = 13.347$, $p < .001$, $\eta_p^2 = .170$) and device ($F(1,130) = 31.384$, $p < .001$, $\eta_p^2 = .194$) for actor within medium. Post-hoc revealed significant differences in the following pairs of Non and Near ($p < .001$), and Semi and Near ($p < .001$): Participants working with the Near-transparent avatar perceived less sensation on their partner than other two avatar conditions. There was no significant main interaction effect between the two factors ($F(2,130) = 1.841$, $p = .163$, $\eta_p^2 = .028$).

A significant main effect was found for device on engagement ($F(1,130) = 28.560$, $p < .001$, $\eta_p^2 = .180$). However, no significant effect was found for transparency ($F(2,130) = 1.911$, $p = .152$, $\eta_p^2 = .029$), nor any interaction effect between the factors ($F(2,130) = .940$, $p = .393$, $\eta_p^2 = .014$).

Nowak's Social Presence (NSP): The two subfactors of NSP—copresence and telepresence—were not aggregated because they consisted of different Likert scale items. The internal consistency of both subfactors was satisfied with the accepted level of Cronbach's alpha (copresence: $\alpha = .890$; telepresence: $\alpha = .870$).

For the copresence, we found significant main effects for both transparency ($F(2,130) = 4.148$, $p = .018$, $\eta_p^2 = .060$) and device ($F(1,130) = 11.518$, $p < .001$, $\eta_p^2 = .081$). In the post-hoc analysis, a significant difference was only found between Semi and Near conditions ($p = .016$): Participants collaborating with the Semi-transparent felt more intimacy and involvement with their partner compared to the Near-transparent avatar. We also found no interaction effect between factors ($F(2,130) = 2.360$, $p = .098$, $\eta_p^2 = .035$).

The telepresence subscale showed a significant main effect only for

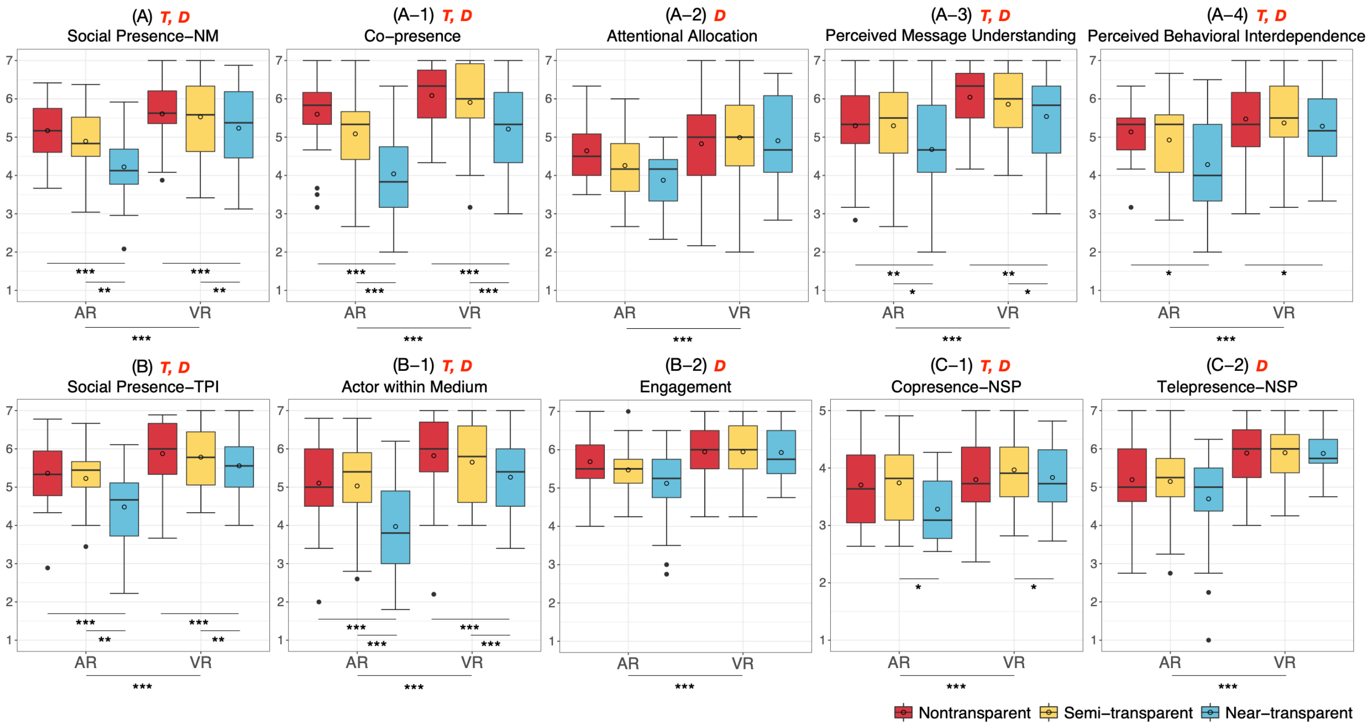

Fig. 3: (A) Social Presence (NM); (A-1) Co-presence; (A-2) Attentional Allocation; (A-3) Perceived Message Understanding; (A-4) Perceived Behavioral Interdependence; (B) Social Presence (TPI); (B-1) Actor within Medium; (B-2) Engagement; (C-1) Copresence and (C-2) Telepresence (T and D: significant main effect of transparency and device, respectively)

device ($F(1,130) = 36.931$, $p < .001$, $\eta_p^2 = .221$). No significant main effect of transparency was found ($F(2,130) = .997$, $p = .372$, $\eta_p^2 = .015$), nor any interaction effect between the factors ($F(2,130) = .631$, $p = .533$, $\eta_p^2 = .010$).

Lastly, in all social presence factors, device type had a significant main effect (Fig. 3): VR participants experienced higher social presence than the AR participants.

4.2 Perceived Action Possibility

The perceived action possibility satisfied internal consistency among the Likert items ($\alpha = .815$). A significant main effect of device was found ($F(1,130) = 22.946$, $p < .001$, $\eta_p^2 = .150$). However, there was no significant main effect of transparency ($F(2,130) = 1.934$, $p = .149$, $\eta_p^2 = .029$), nor any interaction effect between two factors ($F(2,130) = 1.641$, $p = .198$, $\eta_p^2 = .025$): VR participants perceived higher action possibilities in the environment than the AR participants.

4.3 Interpersonal Impression

Due to the emotional reaction asking about negative feelings on the encountered avatar, we adopted reverse-coded values for analysis: The higher values indicate more positive responses. It also showed a reliable Cronbach's alpha value ($\alpha = .828$), and there was a significant main effect of device ($F(1,130) = 13.442$, $p < .001$, $\eta_p^2 = .094$). However, we found no significant main effect of avatar transparency ($F(2,130) = 1.234$, $p = .294$, $\eta_p^2 = .019$), nor interaction effect ($F(2,130) = .786$, $p = .458$, $\eta_p^2 = .012$): AR participants felt more positively about their partner's avatar than the VR participants.

For likability, a significant main effect was found for transparency ($F(2,130) = 4.104$, $p = .019$, $\eta_p^2 = .059$), but not for device ($F(1,130) = .595$, $p = .442$, $\eta_p^2 = .005$). Post-hoc revealed a significant difference between Semi and Near conditions ($p = .029$): Participants preferred the Semi-transparent avatar more than the Near-transparent avatar during collaboration. We found a significant interaction effect between the two factors ($F(2,130) = 3.505$, $p = .033$, $\eta_p^2 = .051$), and the post-hoc showed that: The Non-transparent avatar was more preferred than the Near-transparent avatar by participants in AR compared to VR ($p = .035$).

4.4 Task Load

There was a significant main effect for device on task load ($F(1,128.73) = 5.307$, $p = .023$, $\eta_p^2 = .040$). However, there was no significant main effect of transparency ($F(2,127.05) = .615$, $p = .542$, $\eta_p^2 = .010$), nor significant interaction effect between factors ($F(2,127.05) = .377$, $p = .687$, $\eta_p^2 = .006$): AR Participants perceived higher task load than the VR participants.

4.5 Task Completion Time

Since the participants collaborated together in pairs under the MR configuration, the factor of transparency was only used to compare task completion time. The data were normally distributed (Non: $W = .940$, $p = .292$; Semi: $W = .926$, $p = .167$; Near: $W = .952$, $p = .462$), and the assumption of sphericity was also satisfied ($\chi^2(2) = .669$, $p = .716$); A one-way repeated measures ANOVA indicated that there was no significant difference in task completion time for the three transparency conditions ($F(2,34) = .499$, $p = .611$, $\eta_p^2 = .029$).

4.6 Observations and Interview Comments

We observed and recorded the participant's behavior through in-session monitoring and video recordings to uncover traits that can support the statistical results. Since it was observed that frequent close encounters between the two avatars during the task, we additionally analyzed this behavior (Fig. 5): Close encounters were defined as instances in which the two avatars' bodies came into contact. When these encounters lasted longer than a second, we categorized them as persisting close encounters and measured their duration. Resultingly, we conducted a statistical analysis on the number of close encounters (Fig. 5(I-1)) and the duration of close encounters (Fig. 5(I-2)) to investigate significant differences between transparent avatars.

Close Encounters Between Avatars: The data on the total number of close encounters satisfied the assumptions on normality (Non: $W =$

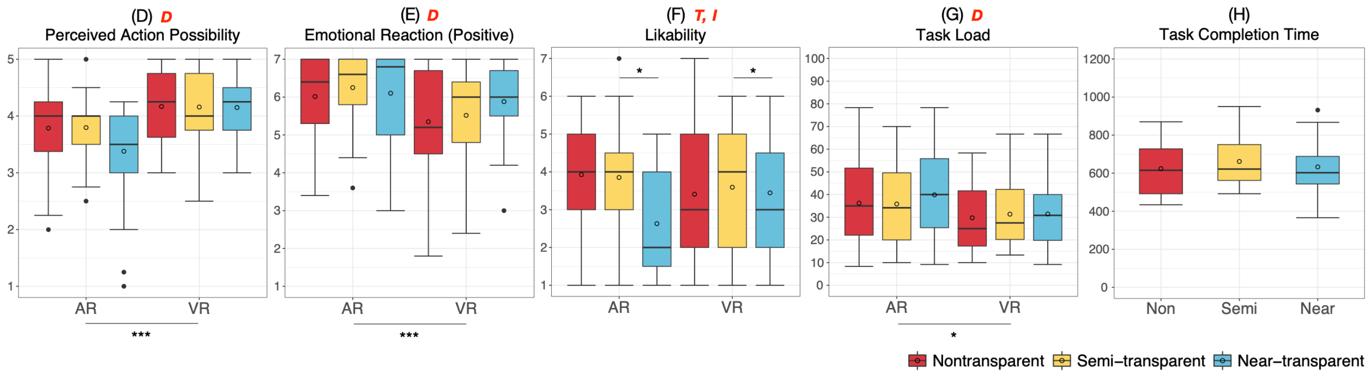

Fig. 4: (D) Perceived Action Possibility; (E) Emotional Reaction; (F) Likability; (G) Task Load; and (H) Task Completion Time (in seconds); (T and D: significant main effect of transparency and device, respectively; I: significant main interaction effect between the two variables)

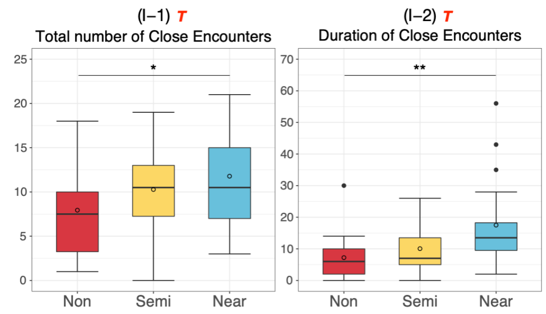

Fig. 5: (I-1) Total number of Close Encounters and (I-2) Duration of Close Encounters (in seconds); (T: significant main effect of transparency)

.935, $p = .241$; Semi: $W = .987$, $p = .995$; Near: $W = .940$, $p = .284$), and homogeneity ($\chi^2(2) = 1.344$, $p = .511$), so we used a one-way repeated measures ANOVA. There was a significant difference in the total number of close encounters on the transparency factor ($F(2,34) = 4.009$, $p = .027$, $\eta_p^2 = .191$), and post-hoc found a significant difference between Non and Near ($p = .043$): Participants were more in contact with Near-transparent avatars than Nontransparent avatars.

Data on the duration of close encounters did not follow normal distribution (Non: $W = .826$, $p = .004$; Semi: $W = .870$, $p = .018$; Near: $W = .844$, $p = .007$). Therefore, the ART ANOVA method was used. There was a significant main effect of transparency ($F(2,34) = 6.938$, $p = .003$, $\eta_p^2 = .290$), and the post-hoc found a significant difference between Non and Near ($p = .002$): Participants were inclined to stay in longer contact with the Near-transparent avatar than the Nontransparent avatar.

Avatar Transparency and Collaboration: To the question of how avatars with different transparency levels affected the overall experience, participants responded similarly, regardless of the device type they wore. First, they mostly mentioned that the transparency of the avatar had no effect when performing individual tasks: They did not feel the difference in the avatar transparency because they mostly focused on the task space, target objects, and their instruction to finish the given task rather than the remote partner (AR-3, VR-8: *"The avatar transparency didn't have a big impact because I didn't have to focus on the avatar when I was doing my task."*). We observed that users were often unaware of their partner while performing individual instructions, thereby attempting to manipulate the same object simultaneously, capturing more in the Near-transparent condition.

Besides, most participants commented that when they had to interact closely with their partner to check what the other was doing or to ask for help, they felt the difference among transparency conditions more keenly: While a highly transparent avatar weakened their partner's presence, a highly opaque avatar made their actions more noticeable (AR-10: *"When the partner was translucent, I could not know where they were so I didn't care. However, when they were opaque, it was easy to see what they needed and were doing, so I could help them."*). Some participants noticed differences when they unexpectedly realized the absence or presence of their partner, for example, by being startled by solely moving objects or an opaque avatar passing by.

Regarding communication and mutual interaction, participants perceived the difference in transparency in situations involving non-verbal expressions such as gestural cues and looking at eyes. However, they responded that on the whole, they could not feel much inconvenience caused by different transparency levels as verbal communication was also possible at all times (VR-27: *"Because I relied mostly on voice, there was no big difference, and the transparency didn't seem to have much effect once communication began."*).

Participants expressed emotional ambivalence towards three transparency conditions. The Nontransparent avatar was found attention-grabbing and reliable, whereas it also caused discomfort and fear due to its strong presence and unrealistic features that became more prominent: They felt sorry when passing through their partner and also wished to avoid causing discomfort. By contrast, the Near-transparent avatar was often considered invisible and unreliable, making it difficult to concentrate on their partner and interpret body gestures. However, it was regarded as less intrusive and allowed for more comfortable interaction, with less concern about collisions and more tolerance of mistakes.

For the Semi-transparent avatar, many participants preferred its visual presence was not too strong nor too weak: This enabled them to maintain a sufficient socio-emotional distance from their partner and reduced the occlusion of visual information in the shared space, whereas this moderate aspect did not leave any lasting impression on some users. (VR-6: *"When the avatar had a proper amount of transparency, I felt less presence but felt less scary."*; AR-24: *"I didn't feel much difference in the middle level."*).

Device Type and Collaboration: The difference in experience between AR and VR users is largely owed to device features and the mediated environment. AR HMD users were difficult to recognize many things at once, due to the limited FoV. The translucent display caused users to doubt the color accuracy of virtual objects and even the Nontransparent avatar was somewhat transparent for AR users, whereas this feature made it easier for them to recognize desired objects behind their partner's avatar. Also, an AR HMD augments virtual objects in real space; the participants preferred to perceive both real and virtual space together because it could relieve anxiety. At the same time, this limited their course of movement in the shared space (AR-26: *"I had to move around to avoid bumping into real furniture, but the VR avatar seemed to be able to move more freely without any restrictions."*). During the collaboration, AR users mostly felt that they needed assistance from their VR partner.

On the contrary, the VR HMD provided a wider FoV, allowing VR users to see more of the shared space and acquire more information at once. As the fully virtual scene was rendered more vividly than in

AR, the VR side was better able to sense their partner's presence in a highly immersive setting; it was frequently observed that VR users often took charge of the collaboration by guiding AR users where to look for certain objects or even bringing them, even if the AR users did not explicitly ask for help. However, many VR users stated that they often felt uncomfortable in situations where their partner blocked their view or ran into one another (VR-2, 17: *"The overlap between avatars when moving or approaching to press the same button made me visually uncomfortable. Moreover, with the opaque avatar, I felt like I might collide with it, or felt more resistance when it passed through me."*). Unlike AR, all elements of the shared space were virtual for the VR users, so they were more inclined to pass through the virtual objects that stood in their way.

5 Discussion

5.1 Analysis on the Study Results

For social presence, the Near-transparent reported lower values on the factors indicating awareness and understanding of the other (related to **R2**), perceived behavioral interdependence (**R3**), and intimacy (**R4**). However, there was no significant difference in factors related to collaboration task involvement (**R1**): We partially accepted H1-1, which assumed that social presence will be lower in the Near-transparent.

Co-presence [16] and actor within medium [30] are concerned with the awareness and sensation of the interactant, and perceived message understanding [16] is related to how well the user understands the interactant's message. As confirmed in the interviews, transparency determined the level of visual information embodied in the remote avatar's image: The more opaque the image, the better a user could recognize and understand the partner. Thus, these factors were rated lower in the most transparent condition than the other two types (**R2**).

Perceived behavioral interdependence [16] indicates how much the user's behavior affects and is affected by the other. This factor showed a significant difference only between the Near-transparent and Nontransparent (**R3**). In the Near-transparent condition, both AR and VR users were unaware of what their partner was doing and failed to respond promptly. We inferred that the weak presence of the Near-transparent avatar may have reduced influences on mutual behavior and awareness, as evidenced by more close encounters and longer persisting duration. Moreover, it also indicates that users did not perceive the Semi-transparent avatar to have affected them in different ways from the most visible and least visible avatars, as it was sufficiently visible yet somewhat transparent at the same time.

Copresence in NSP [36] focuses on intimacy and closeness, as well as involvement in a relationship. Although participants perceived a lower sense of connectedness in the Near-transparent condition (**R4**), they felt socially comfortable interacting with their partners because the dim image made them less conscious of unexpected collisions with the avatar or errors on the other's part. The fact that the Nontransparent avatar did not show statistically higher copresence compared to the other transparent avatars supports this speculation that social comfort and intimacy may be more associated with a less opaque image.

The three factors related to collaboration task and experience involvement showed no significant differences between transparency conditions (**R1**). Attentional allocation [16] is the amount of attention one gives and receives from the partner. Regardless of how visible their partner was, users primarily focused on the task. They paid attention to their partner only when the situation required them to do so; both participants would actively search for one another at the end of a task for a mandatory high-five, as similarly observed by Slater et al. [51]. Therefore, we assume that results are attributed more to what they had to do in the task rather than how visible the avatar was, which was consistent across the conditions.

Similarly, engagement [30] and telepresence [36], measuring the extent of involvement and immersion in the mediated experience, were not affected by transparency. That verbal communication was a given for all conditions may be a reason why the degree of transparency had no effect: They could communicate and focus on what their partner was doing regardless of how well their presence was visually represented. This aligns with Waldow et al. [55] that found no differences on telepresence when users focused on the task space while relying on verbal exchange during collaboration.

We reject H1-2, which hypothesized higher social presence in the AR HMD because all social presence scores were statistically higher for VR than for AR (**R5**). As the AR HMD has a limited FoV and lower resolution than the VR HMD, AR users were not always able to recognize the virtual objects and their partner's avatar as much as the VR users did [34, 49]; however, VR users were able to be more in control and become actively involved, voluntarily taking on a leading role during the task, as it was easier for them to see the shared space and their partner. While our finding is supported by Shin et al. [49] in similar dynamic collaboration settings, this contradicts other studies that found the same factors to be rated higher for AR users [23, 64]. We believe this may be due to these studies not requiring room-scale user movement, where perceiving the entire space becomes more crucial.

Next, we expected that the perceived action possibility [19]—the potential ability to act towards presented objects—will be higher in the Near-transparent condition (H2-1) and the VR HMD (H2-2). A significant effect was found only for the device, with VR users measuring higher than AR users (**R1, R5**): We reject H2-1 and accept H2-2. This owes to the fact that users based their judgment of action possibilities on how the virtual objects, which they could manipulate and were required to do so, were represented [19].

Because VR users were blocked off from reality and fully immersed in the virtual scene, they felt more in control of the virtual object [49]. Conversely, that AR users perceived the virtual space as augmentations on the real environment through a narrower FoV and translucent images limited their command of virtual objects, as physical elements constrained their actions. For both sides, avatar appearance was not a factor that affected them in this regard: While their partner's avatar was also a virtual element in the shared space, it was beyond their control and moved around of its own accord, making it a non-evaluable factor regardless of how visible it was.

Turning to factors of interpersonal impression, we assumed that they would be higher for the Nontransparent (H3-1) and AR HMD (H3-2). Emotional reaction was more positive for AR users than VR (**R5**), but no differences were found between the transparency conditions (**R1**). While AR users preferred the Near-transparent avatar less than the other two types, VR users preferred the Semi-transparent avatar to the Near-transparent avatar (**R4**); We reject H3-1 and partially accept H3-2.

Emotional reaction [4] assessed users' negative feelings, such as anger and surprise, when encountering a virtual counterpart. Contrary to our hypothesis, participants on the whole stated that they regarded the Near-transparent avatar as a "collaborator" co-existing in the shared space somewhere, despite the fact that they could not see it very well. They were not particularly perturbed by its inconspicuous presence, as they had no difficulty in verbally communicating with their partner to complete the task (AR-1: *"I was sometimes surprised during the task, but It didn't cause negative emotions."*). Although Peck et al. [42] found that a more transparent avatar in OST-HMD was perceived as less human-like, their comparison was made for the static image of avatars. As to why VR users perceived their partner's avatar more negatively than AR users, we posit that its relatively stronger presence in the more immersive, vivid environment caused them to feel more uncomfortable when it approached them, as opposed to AR users whose display showed them a less conspicuous version of the avatar for the same transparency [61].

Regarding likability, in line with the copresence, we reconfirm that since the AR OST-HMD rendered the Nontransparent avatar somewhat transparent and less noticeable to begin with, they were perceived as being less invasive than in the VR environment and not much different from the Semi-transparent avatar. In the case of VR, users felt more inclined towards the Semi-transparent avatar because it provided a moderate level of visibility and enabled them to perceive the more vivid environment better.

Lastly, task load had a significant effect on device type, being higher for AR users than VR users (**R5**) but not on avatar transparency (**R1**). Both H4-1 and H4-2, which assumed that the Near-transparent avatar would increase task load and be rated higher by VR users than AR users, were rejected. Although our collaborative task required direct communication between users, its primary goal was for each user to finish individual missions with the help of the other. In the interviews, the majority of participants stated that they were more focused on the task space itself and what was required of them than how they collaborated with their partner. As verbal communication was sufficient to achieve the task, the loss of avatar-contained information due to its transparency became relatively less important.

Contrary to previous studies that found VR induced higher physical and mental efforts [3, 28, 64], AR users felt a higher task load in our study. One possible explanation is that our experimental task demanded more space for the users to explore, search, and manipulate objects, leading them to move around more frequently and see more body images of the remote partner: Compared to VR HMDs, AR OST-HMDs produce less opaque images and have more limited FoVs. Due to these differences in hardware, dynamic changes during the task may have been more burdensome for AR users because they had to pay more effort to grasp the shared space in its entirety within the given device environment.

One notable aspect of the results is that dependent variables that were significantly affected by transparency conditions were not affected by device type: This means that differences between the transparency conditions were perceived in the same way for these variables regardless of the device a user was wearing. Although the AR OST-HMD's display produced a less opaque image for the Nontransparent avatar than the VR HMD did, AR users felt that the Nontransparent avatar was significantly more opaque than the other conditions regarding those factors, just as VR users did. These results are also confirmed by the user's interview.

5.2 Design Implications for Avatar Transparency in Task-centric MR Remote Collaboration

1. For task-centric remote MR collaboration, a strong avatar presence is not necessary: In task-centric MR remote collaboration, the user is mainly interested in the goal, target object, and execution of the task. Therefore, task performance and users' involvement are not much affected by how noticeable the partner's avatar is, particularly when verbal communication is sufficiently provided to aid one another. In such cases, especially those that involve room-scale active movement in a shared virtual space that is densely packed with visual information, raising the level of transparency for the avatars is beneficial: This could help users feel more social comfort and be less conscious about others when they need to focus on what they are doing individually or prevent information related to the task from being frequently occluded.

If users are represented more transparently than other visual elements in the shared space, they can also better acquire task-related information while hiding obtrusive, irrelevant, and unnecessary information about themselves. Transparent avatars can thus be applied to remote collaboration between users who are socially distant or in situations, where there are too many users in the same space such as social VR. Furthermore, utilizing transparent avatars would help hide technical imperfections, save the additional cost for avatar expressions, and focus more on augmenting important content in such task-centric collaboration with limited system resources.

2. Different levels of transparency should be applied to the AR and VR avatars: In terms of different devices, avatars should be presented more opaquely and vividly to the AR user than the VR user. Our study corroborated that different characteristics between AR and VR HMDs lead AR users to experience more difficulties during collaboration, which are mainly concerned with: (1) recognition of virtual elements during collaboration; (2) influences real lighting conditions; (3) constraints imposed by the physical environment; and (4) the quantity and quality of visual information acquired. Therefore, adopting different levels of avatar transparency for each device can help overcome the disadvantages in various aspects during remote collaboration.

(1) Avatars in AR should be represented with relatively low levels of transparency: While the AR environment requires clearer and more solid representations of virtual elements overall, fully opaque avatars and moderately transparent avatars bear no significant difference on the user in task-centric remote MR collaboration where the user is focused more on the task and task-related information than the social aspects of communication with the collaborator. At the same time, near-transparent avatars with a very weak presence should be avoided in the first place because they have a negative impact on the foundation of the task by hindering the ability to refer to important information provided by the collaborator.

(2) Fully opaque avatars should be avoided in VR: Because the VR HMD provides a wide FoV and an opaque, high-resolution display, a non-transparent representation of an avatar in active movement leads to frequent occlusion of virtual information needed in the task. VR users commonly commented that they felt uncomfortable and burdened when they came into contact with a nontransparent avatar. Therefore, raising the level of avatar transparency for the VR side in task-centric MR remote collaboration is recommended. As the Near-transparent condition is less inconspicuous in the VR environment than in the AR, higher degrees of transparency can be applied when the task is prioritized over social interaction with the partner.

5.3 Limitations

There are some limitations that should be investigated in the future. First, it is possible that the avatar transparency was not perceived at the exactly same level by the AR and VR users. While we assign the same alpha value on transparent avatars, visual representation through AR HMD would be affected by the real lighting and appear to be less vivid than how they were shown on the VR side. Due to the AR HMD applying additive blending to display virtual elements, the colors of the avatar and virtual objects may have been represented inaccurately.

Second, more transparency levels should be explored: Because it is largely affected by environmental settings, different alpha ranges should be set as experimental conditions to account for the influence of the virtual environment on the perception of avatar transparency. While we attempted to control these external effects to the best of our ability, investigating various thresholds under different settings will be needed.

Lastly, other collaboration scenarios with different contexts and purposes should also be investigated and compared with our current scenario. While we began our study under the assumption that applying different levels of transparency to the remote avatar would be beneficial for dynamic task-centric MR remote collaboration, how this would affect contrasting situations should also be verified. Furthermore, since many studies mentioned that transparency would have more advantages in remote collaboration with many participants, the scale of users should also be subject to expand the findings of this study.

6 Conclusion

We conclude that applying different transparency levels to avatars in MR remote collaboration should consider the context, such as collaboration purpose and device settings: In task-centric collaboration within a shared space packed with various task-related information, raising the level of transparency is beneficial for users in terms of acquiring the information they need without being overwhelmed by the presence of their partner. Regarding device type, display features of each device and the virtual environment as mediated by them should be considered in determining the level of transparency: While more vivid avatars are suitable for AR HMDs, more transparent and thus less intrusive avatars are beneficial for VR users.

In the future, we will expand the scope of our study to explore transparency effects on different collaboration contexts such as user-centric tasks that involve larger groups of people. Based on the implications of our study, we will also investigate MR collaboration between remote avatars with different transparency levels for each side.

Acknowledgments

This work was supported by the National Research Council of Science & Technology (NST) grant by the Korea government (MSIT) (No. CRC21011) and Institute of Information & communications Technology Planning Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2019-0-01270, WISE AR UI/UX Platform Development for Smartglasses)

REFERENCES

[1] F. Argelaguet, A. Kulik, A. Kunert, C. Andujar, and B. Froehlich. See-through techniques for referential awareness in collaborative virtual reality. *International Journal of Human-Computer Studies*, 69(6):387–400, 2011. 1

[2] S. Aseeri and V. Interrante. The influence of avatar representation on interpersonal communication in virtual social environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2608–2617, 2021. 1, 2

[3] H. Bai, P. Sasikumar, J. Yang, and M. Billinghurst. A user study on mixed reality remote collaboration with eye gaze and hand gesture sharing. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pp. 1–13, 2020. 2, 9

[4] J. N. Bailenson, J. Blascovich, A. C. Beall, and J. M. Loomis. Interpersonal distance in immersive virtual environments. *Personality and social psychology bulletin*, 29(7):819–833, 2003. 4, 8

[5] B. Benda and E. D. Ragan. The effects of virtual avatar visibility on pointing interpretation by observers in 3d environments. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 50–59. IEEE, 2021. 2

[6] S. Benford, C. Greenhalgh, G. Reynard, C. Brown, and B. Koleva. Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on computer-human interaction (TOCHI)*, 5(3):185–223, 1998. 2

[7] F. Biocca, C. Harms, and J. K. Burgoon. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & virtual environments*, 12(5):456–480, 2003. 2

[8] Y. Cao, X. Qian, T. Wang, R. Lee, K. Huo, and K. Ramani. An exploratory study of augmented reality presence for tutoring machine tasks. In *Proceedings of the 2020 CHI conference on human factors in computing systems*, pp. 1–13, 2020. 1

[9] Y. F. Cheng, H. Yin, Y. Yan, J. Gugenheimer, and D. Lindlbauer. Towards understanding diminished reality. In *CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2022. 1

[10] S. Cho, S.-w. Kim, J. Lee, J. Ahn, and J. Han. Effects of volumetric capture avatars on social presence in immersive virtual environments. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 26–34. IEEE, 2020. 1

[11] C. Dubosc, G. Gorisse, O. Christmann, S. Fleury, K. Poinsot, and S. Richir. Impact of avatar anthropomorphism and task type on social presence in immersive collaborative virtual environments. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 438–439. IEEE, 2021. 1

[12] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsomboon, and M. Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81–98, 2019. 2

[13] G. Gamelin, A. Chellali, S. Cheikh, A. Ricca, C. Dumas, and S. Otmane. Point-cloud avatars to improve spatial communication in immersive collaborative virtual environments. *Personal and Ubiquitous Computing*, 25(3):467–484, 2021. 1, 2

[14] J. G. Grandi, H. G. Debarba, and A. Maciel. Characterizing asymmetric collaborative interactions in virtual and augmented realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 127–135. IEEE, 2019. 2

[15] S. W. Greenwald, Z. Wang, M. Funk, and P. Maes. Investigating social presence and communication with embodied avatars in room-scale virtual reality. In *International Conference on Immersive Learning*, pp. 75–90. Springer, 2017. 2

[16] C. Harms and F. Biocca. Internal consistency and reliability of the networked minds measure of social presence. 2004. 4, 8

[17] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, pp. 904–908. Sage publications Sage CA: Los Angeles, CA, 2006. 4

[18] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. In *Advances in psychology*, vol. 52, pp. 139–183. Elsevier, 1988. 4

[19] T. Hartmann, W. Wirth, H. Schramm, C. Klimmt, P. Vorderer, A. Gysbers, S. Böcking, N. Ravaja, J. Laarni, T. Saari, et al. The spatial presence experience scale (spes): A short self-report measure for diverse media settings. *Journal of Media Psychology: Theories, Methods, and Applications*, 28(1):1, 2016. 4, 8

[20] Z. He, R. Du, and K. Perlin. Collabovr: A reconfigurable framework for creative collaboration in virtual reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 542–554. IEEE, 2020. 1

[21] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social vr. In *2017 IEEE symposium on 3D user interfaces (3DUI)*, pp. 233–234. IEEE, 2017. 2

[22] W. A. IJsselsteijn, H. De Ridder, J. Freeman, and S. E. Avons. Presence: concept, determinants, and measurement. In *Human vision and electronic imaging V*, vol. 3959, pp. 520–529. SPIE, 2000. 2

[23] D. Jo, K.-H. Kim, and G. J. Kim. Effects of avatar and background types on users' co-presence and trust for mixed reality-based teleconference systems. In *Proceedings the 30th Conference on Computer Animation and Social Agents*, pp. 27–36, 2017. 1, 2, 8

[24] S. Kim, G. Lee, W. Huang, H. Kim, W. Woo, and M. Billinghurst. Evaluating the combination of visual communication cues for hmd-based mixed reality remote collaboration. In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–13, 2019. 4

[25] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pp. 3–12. IEEE, 2010. 1, 2

[26] E. Langbehn, E. Harting, and F. Steinicke. Shadow-avatars: A visualization method to avoid collisions of physically co-located users in room-scale vr. In *IEEE Workshop on Everyday Virtual Reality*, 2018. 2

[27] C. Lee, G. A. Rincon, G. Meyer, T. Höllerer, and D. A. Bowman. The effects of visual realism on search tasks in mixed reality simulation. *IEEE transactions on visualization and computer graphics*, 19(4):547–556, 2013. 1

[28] G. A. Lee, T. Teo, S. Kim, and M. Billinghurst. A user study on mr remote collaboration using live 360 video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 153–164. IEEE, 2018. 9

[29] Y. Li, E. Ch'ng, S. Cobb, and S. See. Presence and communication in hybrid virtual and augmented reality environments. *PRESENCE: Virtual and Augmented Reality*, pp. 1–40, 2021. 2

[30] M. Lombard, T. B. Ditton, and L. Weinstein. Measuring presence: the temple presence inventory. In *Proceedings of the 12th annual international workshop on presence*, pp. 1–15, 2009. 4, 8

[31] M. Martini, K. Kilteni, A. Maselli, and M. V. Sanchez-Vives. The body fades away: investigating the effects of transparency of an embodied virtual body on pain threshold and body ownership. *Scientific reports*, 5(1):1–8, 2015. 2, 3

[32] M. Matamala-Gomez, A. M. D. Gonzalez, M. Slater, and M. V. Sanchez-Vives. Decreasing pain ratings in chronic arm pain through changing a virtual body: different strategies for different pain types. *The Journal of Pain*, 20(6):685–697, 2019. 3

[33] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994. 2

[34] M. R. Miller and J. N. Bailenson. Social presence outside the augmented reality field of view. *Frontiers in Virtual Reality*, 2:656473, 2021. 8

[35] M. Norman, G. A. Lee, R. T. Smith, and M. Billingurst. The impact of remote user's role in a mixed reality mixed presence system. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry*, pp. 1–9, 2019. 1, 4

[36] K. L. Nowak and F. Biocca. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 12(5):481–494, 2003. 4, 8

[37] K. L. Nowak and J. Fox. Avatars and computer-mediated communication: A review of the definitions, uses, and effects of digital representations. *Review of Communication Research*, 6:30–53, 2018. 1

[38] S. Orts-Escolano, C. Rhemann, S. Fanello, W. Chang, A. Kowdle, Y. Degtyarev, D. Kim, P. L. Davidson, S. Khamis, M. Dou, et al. Holoportation: Virtual 3d teleportation in real-time. In *Proceedings of the 29th annual symposium on user interface software and technology*, pp. 741–754, 2016. 1

[39] N. Osmers, M. Prilla, O. Blunk, G. George Brown, M. Janßen, and N. Kahrl. The role of social presence for cooperation in augmented reality on head mounted devices: A literature review. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2021. 4

[40] M. Pakanen, P. Alavesa, N. van Berkel, T. Koskela, and T. Ojala. "nice to see you virtually": Thoughtful design and evaluation of virtual avatar

of the other user in ar and vr based telexistence systems. *Entertainment Computing*, 40:100457, 2022. 2

[41] Y. Pan and A. Steed. The impact of self-avatars on trust and collaboration in shared virtual environments. *PloS one*, 12(12):e0189078, 2017. 1

[42] T. C. Peck, J. J. Good, A. Erickson, I. Bynum, and G. Bruder. Effects of transparency on perceived humanness: Implications for rendering skin tones using optical see-through displays. *IEEE Transactions on Visualization and Computer Graphics*, 28(5):2179–2189, 2022. 2, 3, 8

[43] T. Piumsomboon, A. Dey, B. Ens, G. Lee, and M. Billinghurst. The effects of sharing awareness cues in collaborative mixed reality. *Frontiers in Robotics and AI*, 6:5, 2019. 2

[44] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–13, 2018. 1, 2

[45] H. Regenbrecht, K. Meng, A. Reepen, S. Beck, and T. Langlotz. Mixed voxel reality: Presence and embodiment in low fidelity, visually coherent, mixed reality environments. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 90–99. IEEE, 2017. 2

[46] J. Reinhardt and K. Wolf. Go-through: Disabling collision to access obstructed paths and open occluded views in social vr. In *Proceedings of the Augmented Humans International Conference*, pp. 1–10, 2020. 1, 2

[47] T. Rhee, S. Thompson, D. Medeiros, R. Dos Anjos, and A. Chalmers. Augmented virtual teleportation for high-fidelity telecollaboration. *IEEE transactions on visualization and computer graphics*, 26(5):1923–1933, 2020. 1, 2

[48] A. Schäfer, G. Reis, and D. Stricker. A survey on synchronous augmented, virtual, andmixed reality remote collaboration systems. *ACM Computing Surveys*, 55(6):1–27, 2022. 4

[49] J.-E. Shin, B. Yoon, D. Kim, H.-I. Kim, and W. Woo. The effects of device and spatial layout on social presence during a dynamic remote collaboration task in mixed reality. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 394–403. IEEE, 2022. 1, 2, 8

[50] K.-s. Shin, H. Kim, J. g. Lee, and D. Jo. Exploring the effects of scale and color differences on users' perception for everyday mixed reality (mr) experience: Toward comparative analysis using mr devices. *Electronics*, 9(10):1623, 2020. 1, 2

[51] M. Slater, A. Sadagic, M. Usoh, and R. Schroeder. Small-group behavior in a virtual and real environment: A comparative study. *Presence*, 9(1):37–51, 2000. 8

[52] M. Speicher, B. D. Hall, and M. Nebeling. What is mixed reality? In *Proceedings of the 2019 CHI conference on human factors in computing systems*, pp. 1–15, 2019. 2

[53] A. Tang, F. Biocca, L. Lim, et al. Comparing differences in presence during social interaction in augmented reality versus virtual reality environments: An exploratory study. *Proceedings of PRESENCE*, pp. 204–208, 2004. 1

[54] T. Teo, M. Norman, G. A. Lee, M. Billinghurst, and M. Adcock. Exploring interaction techniques for 360 panoramas inside a 3d reconstructed scene for mixed reality remote collaboration. *Journal on Multimodal User Interfaces*, 14(4):373–385, 2020. 2

[55] K. Waldow, A. Fuhrmann, and S. M. Grünvogel. Investigating the effect of embodied visualization in remote collaborative augmented reality. In *International Conference on Virtual Reality and Augmented Reality*, pp. 246–262. Springer, 2019. 1, 2, 8

[56] F. Weidner, G. Boettcher, S. A. Arboleda, C. Diao, L. Sinani, C. Kunert, C. Gerhardt, W. Broll, and A. Raake. A systematic review on the visualization of avatars and agents in ar & vr displayed using head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 2023. 1, 4

[57] T. Weissker and B. Froehlich. Group navigation for guided tours in distributed virtual environments. *IEEE Transactions on Visualization and Computer Graphics*, 27(5):2524–2534, 2021. 2

[58] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 143–146, 2011. 5

[59] A. E. Wolf, J. Reinhardt, M. Kurzweg, and K. Wolf. Socialslider: Changing the transparency of avatars. In *Augmented Humans 2022*, pp. 276–283. 2022. 1, 2

[60] E. Wolf, M. L. Fiedler, N. Döllinger, C. Wienrich, and M. E. Latoschik. Exploring presence, avatar embodiment, and body perception with a holographic augmented reality mirror. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 350–359. IEEE, 2022. 2

[61] E. Wolf, D. Mal, V. Frohnapfel, N. Döllinger, S. Wenninger, M. Botsch, M. E. Latoschik, and C. Wienrich. Plausibility and perception of personalized virtual humans between virtual and augmented reality. In *2022 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 489–498. IEEE, 2022. 2, 8

[62] Y. Wu, Y. Wang, S. Jung, S. Hoermann, and R. W. Lindeman. Using a fully expressive avatar to collaborate in virtual reality: Evaluation of task performance, presence, and attraction. *Frontiers in Virtual Reality*, 2:641296, 2021. 1, 2

[63] B. Yoon, H.-i. Kim, G. A. Lee, M. Billinghurst, and W. Woo. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 547–556. IEEE, 2019. 1, 2, 4

[64] B. Yoon, H.-i. Kim, S. Y. Oh, and W. Woo. Evaluating remote virtual hands models on social presence in hand-based 3d remote collaboration. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 520–532. IEEE, 2020. 1, 2, 4, 8, 9

[65] K. Yu, G. Gorbachev, U. Eck, F. Pankratz, N. Navab, and D. Roth. Avatars for teleconsultation: effects of avatar embodiment techniques on user perception in 3d asymmetric telepresence. *IEEE Transactions on Visualization and Computer Graphics*, 27(11):4129–4139, 2021. 1, 2, 4