Mirror Me: Exploring Avatar Self-Views in Multi-user Virtual Reality

KATJA KRUG[*†], Interactive Media Lab Dresden, Dresden University of Technology, Germany
DANIEL IMMANUEL FINK[*‡], HCI Group, University of Konstanz, Germany
MATS OLE ELLENBERG[†], Interactive Media Lab Dresden, Dresden University of Technology, Germany
ANKE V. REINSCHLUESSEL[‡], HCI Group, University of Konstanz, Germany
WOLFGANG BÜSCHEL, VISUS, University of Stuttgart, Germany
TIARE FEUCHTNER[‡], HCI Group, University of Konstanz, Germany
RAIMUND DACHSELT[†§], Interactive Media Lab Dresden, Dresden University of Technology, Germany

Fig. 1. In our study, we explore three types of self-views in VR: A Remote Perspective View (B), a Personal Mirror (C), and a Miniature Avatar (D) in comparison to a baseline with No Additional Self-View (A). The self-views are visualized with a purple border in context with the remote avatar with a blue border.

This paper investigates avatar self-views in virtual reality (VR) and their impact on multi-user communication scenarios. Self-views in video conferencing have been shown to impact our self-perception and mental load, so we explore whether similar effects occur in VR, as personal and professional gatherings progressively move to virtual spaces with 3D avatars. We identify the following key design dimensions for self-representations in VR: spatiality, anchoring and size, and self-visibility. Based on these, we designed three variants (Remote Perspective View, Personal Mirror, and Miniature Avatar), which we compare to a baseline (No Additional Self-View) in a user study. Our analysis of sixteen dyads playing a word-guessing game requiring verbal and non-verbal communication (i.e., explaining and charades) in VR confirms that self-views are

[*]The first two authors contributed equally to this research.
[†]Katja Krug, Mats Ole Ellenberg, and Raimund Dachselt are also with the Centre for Tactile Internet with Human-in-the-Loop (CeTI)
[‡]Daniel Immanuel Fink, Anke V. Reinschluessel, and Tiare Feuchtner also with the Centre for the Advanced Study of Collective Behaviour (CASCB)
[§]Raimund Dachselt is also with the Centre for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI)

Authors' Contact Information: Katja Krug, Interactive Media Lab Dresden, Dresden University of Technology, Dresden, Germany, katjakrug@acm.org; Daniel Immanuel Fink, HCI Group, University of Konstanz, Konstanz, Germany, daniel.immanuel.fink@uni-konstanz.de; Mats Ole Ellenberg, Interactive Media Lab Dresden, Dresden University of Technology, Dresden, Germany, mellenberg@acm.org; Anke V. Reinschluessel, HCI Group, University of Konstanz, Konstanz, Germany, anke.reinschluessel@uni-konstanz.de; Wolfgang Büschel, VISUS, University of Stuttgart, Stuttgart, Germany, bueschel@acm.org; Tiare Feuchtner, HCI Group, University of Konstanz, Konstanz, Germany, tiare.feuchtner@uni-konstanz.de; Raimund Dachselt, Interactive Media Lab Dresden, Dresden University of Technology, Dresden, Germany, dachselt@acm.org.

ACM 2573-0142/2025/12-ARTISS001
https://doi.org/10.1145/3773058

beneficial for communication scenarios that require expressive body language or facial expressions, as this allows monitoring the own avatar's animations. Further, the Miniature Avatar was overall preferred due to its spatiality, world-anchoring, and full self-visibility. Based on our results, we offer design recommendations for self-views covering the key design dimensions in VR.

CCS Concepts: • **Human-centered computing → Empirical studies in HCI**; **Virtual reality**.

Additional Key Words and Phrases: virtual reality meetings, video conferencing, empirical study, self - representation, virtual reality, dyad setting, multi-user, VR

ACM Reference Format:
Katja Krug, Daniel Immanuel Fink, Mats Ole Ellenberg, Anke V. Reinschluessel, Wolfgang Büschel, Tiare Feuchtner, and Raimund Dachselt. 2025. Mirror Me: Exploring Avatar Self-Views in Multi-user Virtual Reality. *Proc. ACM Hum.-Comput. Interact.* 9, 8, Article ISS001 (December 2025), 26 pages. https://doi.org/10.1145/3773058

1 Introduction

Whether it is catching up with loved ones or joining a team meeting, many people increasingly rely on video conferencing platforms like Zoom[1] or WhatsApp Video[2] to stay connected in a physically distant world. Unlike face-to-face interactions, a unique feature of these platforms is the ability to constantly monitor ourselves through a self-view, i.e., our own camera image displayed alongside our conversation partners. This gives us insights into how we are perceived, enables us to take control over what we wish to share with others, to adjust our appearance, and to ensure the perceivability of our non-verbal communication.

As social Virtual Reality (VR) gains traction, remote meetings are speculated to gradually shift towards immersive virtual spaces, i.e., the Metaverse [12], where users wear head-mounted displays (HMDs) to interact with each other via 3D avatars instead of live video feeds. So, instead of a real-time reflection of themselves, users are represented by a potentially stylized virtual avatar. Therefore, we want to explore how the well-established concept of self-views during video conferencing translates to social VR spaces. The appearance of user representations [54] has been shown to be impactful in social VR settings, with prior research examining different forms of avatar representations in multi-user settings [21], their effect on social and co-presence [31, 33, 70], user-avatar relationship [32], and communication efficiency and quality [4, 22, 23]. Further, the unfamiliarity with the looks of an avatar and the mismatch between one's real life appearance and virtual representation makes VR self-views potentially influential during conversation. Users also mostly lack awareness how tracked facial or body expressions are reproduced onto their own avatars, creating a need for reassurance through self-monitoring and a desire to better understand their non-verbal contribution to the conversation. However, this could make the interaction susceptible to fatiguing effects similar to "Zoom Fatigue", a phenomenon where heightened self-awareness in video conferencing scenarios contributes to exhaustion [6, 57]. The dual role of self-views in video conferences, as both a tool for control and a potential source of fatigue, raises intriguing questions about their implications in immersive virtual environments. Therefore, we aim to explore whether users desire additional self-views in VR, explore their design space, and investigate their potential impact on the experience, as well as opportunities and limitations.

Inspired by video conferencing tools, we conceptualized multiple VR self-view variants covering a range of design possibilities within our design space. We chose the following three variants, differing in the key dimensions *spatiality, anchoring and size, and self-visibility* of our design space, for this explorative study: (1) A Remote Perspective View, showing the user from their

[1]Zoom: https://www.zoom.com/ (last accessed: October 27th, 2025)
[2]WhatsApp video calls: https://www.whatsapp.com/calling (last accessed: October 27th, 2025)

conversation partner's perspective (see Figure 1B), (2) a Personal Mirror (see Figure 1C), and (3) a Miniature Avatar (see Figure 1D). In a user study with 16 dyads (32 participants), as an initial exploration into the smallest multi-user setting, we compared these VR self-views to a baseline with no additional self-view (see Figure 1A). The dyads played a word-guessing game in VR involving both (1) verbal explanations in a conversation-focused setting and (2) gesture-based communication, including body and hand gestures and facial expressions (i.e., charades). We captured the users' gaze behavior and inquired about their attention distribution, self-assurance, and perceived helpfulness of the self-view. To investigate potential fatiguing effects, we also measured their perceived mental workload. We combine the resulting data with insights into general preferences from interviews after the experiment.

Our findings highlight the benefit of self-views in dual-user communication scenarios that require expressive body language. While Personal Mirror was the least preferred one and viewed as distracting due to its head-coupled nature, and Remote Perspective View drew most of the attention away from the partner, the Miniature Avatar was the most preferred option due to its flexible placement and its helpfulness in providing an overview of all the avatar's movements.

In summary, our main contributions are:

- An exploratory analysis of three self-view variants in VR regarding perceived helpfulness, mental load, impact on self-confidence, and attention distribution.
- Insights into user preferences regarding the use and design of these self-views.
- Design recommendations for VR self-views based on our measurements and user feedback.

2 Related Work

With the rise of social VR platforms ranging from casual (e.g., Meta Horizon Worlds[3], VRChat[4]) to professional work contexts (e.g., Meta Horizon Workrooms[5], Microsoft Mesh[6]), immersive spaces are often speculated to be the future of online meetings. In VR spaces, people embody virtual avatars, which can enhance social presence and enable conversation patterns similar to face-to-face interactions [58]. While the quality of the avatar's appearance and non-verbal expressiveness affects communication quality, the anonymity they provide can also make people more at ease [10, 64]. In contrast, video conferences have been linked to heightened exhaustion compared to in-person meetings [6]. The direct comparison between VR and video conferencing highlights key differences in communication behavior: video conferences elicit increased verbal and non-verbal back-channeling, and VR interactions rely more on deictic gestures and exhibit gaze patterns closer to in-person communication [2]. This further positions multi-user VR as the missing link between video conferences and in-person meetings. However, existing research has not yet explored whether self-views – a familiar concept from video conferencing – are also beneficial in VR, or if the resemblance to in-person meetings and the use of avatars make self-observation unnecessary. Therefore, we will first look into the role of self-views in video conferencing and the impact of the appearance and visibility of our self-representations, i.e., avatars, in VR.

2.1 The Impact of Self-Views in Video Conferencing

During the COVID-19 pandemic, video conferencing became essential for maintaining personal and professional connections. However, many reported increased fatigue in comparison to face-to-face meetings, impacting their cognitive performance [44], among other things. Factors such as awkward turn-taking and inhibited spontaneity contribute to this exhaustion [1]. Bailenson [6] identified four

[3]Meta Horizon Worlds: https://horizon.meta.com/ (last accessed: May 30th, 2025)
[4]VRChat: https://vrchat.com/ (last accessed: May 30th, 2025)
[5]Meta Horizon Workrooms: https://forwork.meta.com/horizon-workrooms/ (last accessed: May 30th, 2025)
[6]Microsoft Mesh: https://microsoft.com/microsoft-teams/microsoft-mesh (last accessed: May 30th, 2025)

main causes of this *Zoom Fatigue*: (1) Restricted mobility in the camera frame, (2) excessive, close-up "eye contact", (3) increased cognitive load from actively managing own non-verbal cues, and (4) the constant self-view fostering self-consciousness and self-criticism. Out of these aspects, the efforts going into self-representation as a major contributor to this fatigue particularly affect female and newer employees, due to societal pressures on appearance and competence [57]. Additionally, factors like "mirror anxiety" and the hyper-awareness of being watched [18] add to the burden of female users.

This negative self-perception might be reduced through habituation, as it has been shown that users become more comfortable with their appearance over time [27]. Accordingly, despite the cognitive strain, most users keep their self-view enabled for self-monitoring, often unaware of the option to disable it, and they preferably use the gallery option, where everyone's video window is the same size [7]. Miller et al. [41] further suggest that self-views can reduce the effects of social anxiety.

Many video conferencing challenges may be mitigated in VR, such as offering increased mobility and free movement within a larger tracking area. Some of the challenges, such as the mental strain of having all participants face you at all time [6] without actually making true eye contact [1], have already been addressed in video-based 3D solutions [29, 45, 56, 59, 61], such as the virtual meeting setup introduced by Schuessler et al. [52], which captures users with several cameras and displays them to their spatially-arranged partners from the correct perspective by synthesizing head rotations. VR environments equipped with eye, face, and full-body tracking can further improve gaze alignment and non-verbal communication, allowing users to make comfortable and context-appropriate eye contact. Lastly, customizable 3D avatars could alleviate some of the appearance-related pressures associated with video conferencing.

In summary, while self-views in video conferencing can contribute to fatigue, they remain a preferred feature for self-monitoring and might benefit from habituation effects. As VR brings several advantages to virtual meetings that may reduce the negative effects associated with self-views, we see potential to explore their design and impact in immersive environments.

2.2 The Impact of Self-Representation in VR

Avatars as a self-representation are central to how we experience social gatherings in VR. We are not only affected by the appearance of other people's avatars, but also by our own, which is why we care deeply about how our representation looks. Customizing and personalizing avatars enhances embodiment, body ownership, presence [62], as well as self-identification [20, 48], with matching gender and ethnicity playing key roles [14]. Users maintain a strong bond and identify with their personalized avatars even after swapping bodies [16], and many prefer avatars that reflect nuanced personal traits, even if those are commonly deemed undesirable [43]. Knowing that our avatar represents us in a personalized way can heighten our excitement and satisfaction during a conversation [49] and fosters overall emotionally more positive collaborative experiences, especially for extroverted individuals [11]. These findings suggest that how we are represented in multi-user virtual spaces is important and impactful to us, even if we don't see ourselves, leading to the assumption that a self-view of our embodied avatar will have an effect on our interactions.

Beyond appearance, avatars also subconsciously shape our behavior through the so-called *Proteus effect*, which is "the hypothesis that an individual's behavior conforms to their digital self-representation independent of how others perceive them" [67]. The perceived attractiveness of our avatars influences our behavior in VR spaces, for example, taller avatars encourage more aggressive negotiation [69], especially when digital embodiment is involved [68]. Viewing one's avatar in a virtual mirror can even induce physical sensations, such as seated users feeling sensations of walking leg movements [40]. This effect occurs even though the virtual mirror does not reflect

the user's real physical body, in contrast to mixed reality mirrors [71, 72], which augment the real body and thereby foster a natural connection with the reflection. However, self-views in VR might also have drawbacks. Döllinger et al. [17] found that seeing your own VR self-representation in a mirror affects your body awareness, i.e., users noticed fewer signals from within their bodies and relied more on visual feedback than physical sensations.

Most of the aforementioned studies have focused on single-user contexts, and those in multi-user contexts only enable self-views before the collaboration, and not throughout. We, therefore, want to explore how continuously visible self-views affect users in multi-user social VR. The research conducted by Ma and Pan [39] might be an indication of a possible outcome: They found that users engaged in a mentally demanding task paid less attention to their virtual mirror reflection, suggesting that task complexity may significantly influence the effects of self-views. Thus, the task in a multi-user scenario must be carefully considered.

3 Self-View, Avatar, and Environment Design

For the design and exploration of avatar self-views in VR, we first identified three key design dimensions for self-representations in both non-immersive and immersive applications from related work and communication applications (see section 3.1). We designed a variety of different self-view variants covering our design dimensions. To assess the self-views, we built a prototype that allows multiple users to meet in VR while being represented by an avatar that they customize beforehand (see section 3.3). Based on pilot tests, we carefully chose and identified three variants covering different aspects (see section 3.2). We extended our prototype to place the users into a virtual furnished room and briefly go into detail on the environment and avatar design in section 3.3. The implementation details are summarized in section 3.4 to support reproducibility efforts.

3.1 Design Dimensions for Self-Views

Based on previous work and communication applications that incorporate user self-representations, we identified the following three key design dimensions for self-views: *Spatiality*, *Anchoring and Size*, and *Self-Visibility*.

Spatiality. Self-views can be categorized on a spatiality continuum ranging from *low* (i.e., "flat" self-views) to *high* (i.e., fully spatial self-views). Most video conferencing applications incorporate flat self-views with *low spatiality* by integrating the user's own camera image into the 2D user interface. In VR, several applications integrate virtual mirrors as self-views (e.g., Meta Horizon Home, [25, 62, 65]), which are basically flat but provide *increased spatiality*, such as depth cues, depending on the implementation. However, VR environments also allow for self-views with *high spatiality*, such as representing the user as a 3D miniature avatar [9, 47].

Anchoring and Size. The anchoring and size of self-views can vary depending on the application and the selected mode. The gallery view, commonly integrated into many video conferencing applications (e.g., Zoom), presents the user's own camera image next to the remote users' images with the *same size* and importance. However, several existing applications also support other modes that show the self-view *bigger* while speaking and *smaller* at other times (speaker view) – or even permanently *small* (e.g., Microsoft Teams). The spatiality of VR environments further expands the placement possibilities for objects like self-views, allowing them to be *world-anchored* (e.g., virtual mirrors), *head- or body-anchored* to the local user, or *object-anchored*, such as being anchored to the remote user's representation [37].

Self-Visibility. The extent to which the own body and movements are visible in the self-view depends on both the design and the employed hardware. In video conferencing applications, self-visibility is mainly determined by the camera's position and perspective defining its field of view (FoV), which commonly covers the *head and shoulders*. However, VR applications usually track the position of various upper body parts using multiple cameras and sensors integrated into the VR device and map them onto a user representation, such as a 3D avatar. The 3D representation of the user in the VR environment allows the design of self-views simulating cameras or mirrors with defined or dynamically adjusted perspectives, as well as self-views like 3D miniature avatars that are not limited to only one perspective, covering the *upper body* or even the *full body*.

3.2 Self-Views in VR

Covering the identified design dimensions, we initially designed and implemented two self-view variants in VR: A flat self-view with *low spatiality* and a fully spatial self-view with *high spatiality*. Both could be *anchored* to the *world*, the local user's *head*, or the remote user's representation (i.e., *object*), and were adjustable in *size*, thereby influencing *self-visibility*. Our flat self-view acts like a square mirror in the virtual environment, while the spatial self-view is a scaled 3D representation of the user's own avatar, mirroring the user's body movements and facial expressions in real-time. With the design options described, we tested a variety of self-view property distributions in exploratory pilot tests. The tested designs ranged from bigger to smaller self-views anchored in a world position, the local user's head (i.e., perspective), or the remote user's head.

Our design exploration revealed that large mirrors consistently occlude a large area of the virtual environment, so they need to be placed to the side of the user to remain visible and not get in the way. As a result, noticeable movements (e.g., head-turning) become necessary to observe the self-view, potentially interrupting the current task, which is why we excluded large mirror self-views from our comparison. Regarding small mirror designs, we experienced that the visibility of flat self-views is highly dependent on their orientation and the user's perspective. As a result, only flat designs anchored to the local or remote user are practical, as they ensure continuous visibility of the self when viewing the self-view, as well as enable an unobstructed view of the opposite user. In contrast, user-anchored self-views with high spatiality limited the ability to view the self-view from different perspectives, reducing their spatial advantages.

Based on the insights from our exploratory pilot tests, we carefully created three different self-view designs that illustratively represent the spectrum of our identified design dimensions (see section 3.1): The Remote Perspective View, the Personal Mirror, and the Miniature Avatar. The designs cover self-views ranging from low to high spatiality, anchored to either the local user (Personal Mirror), the remote user (Remote Perspective View), or in a world position (Miniature Avatar), as informed by our design exploration. For self-view designs close to the user (Personal Mirror and Miniature Avatar), participants were allowed to reposition and resize them through direct interaction at the beginning, while the distant self-view (Remote Perspective View) remained unchanged. All self-view variants are only visible to the local user. In the following, we describe our self-views and the underlying design considerations in more detail.

3.2.1 No Add. Self-View. This variant only shows the first-person avatar embodied by the user. We use it as a baseline for the study, representing the status quo of social VR applications. It does not include any *additional* self-view (see Figure 1A and 2A).

3.2.2 Remote Perspective View. This self-view with rather low spatiality is anchored to the remote user (object-anchored) and inspired by the gallery view of most video conferencing solutions, adapted to VR. Here, the local self-representation is captured from the perspective of the remote user and displayed on a screen-like surface next to the remote avatar's head (see Figure 1B and 2B).

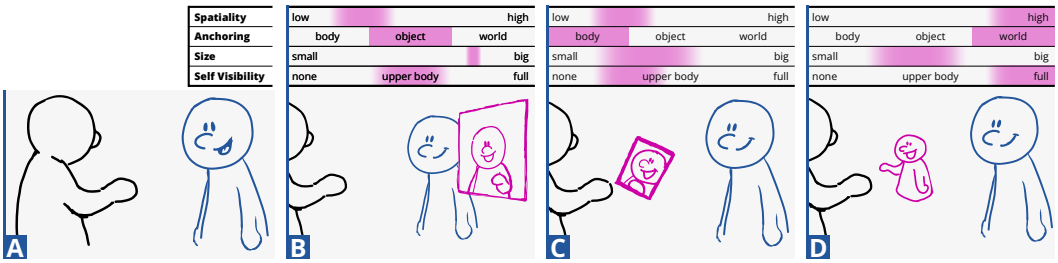


Fig. 2. The categorization of our three VR self-view designs – Remote Perspective View (B), Personal Mirror (C), and Miniature Avatar (D) – based on our identified design dimensions for self-views (see 3.1), in comparison to the baseline No Add. Self-View (A). The self-view is visualized in purple in context with the local user's embodied avatar (black) and the remote avatar (blue).

The virtual camera capturing the local user is positioned at the head of the remote avatar and is always directed towards the head of the local user, regardless of its movement. This way, akin to video conferencing, the local user always sees themselves from the perspective of the remote user, regardless of the remote user's actual viewing direction. The camera frame covers the local user's head, shoulder, and part of their torso, which is a common amount of self-visibility during video calls. The self-view screen is attached to the remote avatar's head position, following its movement, ensuring that the self-view is always displayed next to the remote avatar. The flat screen always rotates toward the local user's head, ensuring continuous visibility (see Figure 3A and 3B). The camera image is mirrored, similar to the default setting of many video conferencing applications. With this method, users can always observe themselves while looking in the direction of the remote user.

3.2.3 Personal Mirror. This is a self-view mirror with rather low spatiality that is anchored to the local user (see Figure 1C and 2C) and follows the user's head rotation around the yaw-axis (see Figure 3C and 3D), but is independent of other head movements (roll and pitch movements). With this behavior, it does not follow the user's FoV, but acts as a body-anchored interface initially positioned at torso height. This is a a common approach for placing a VR interface near the user's body that was shown to be beneficial depending on the content and scenario [8, 38, 55]. To find the most comfortable position, it can be scaled and repositioned relative to the user's body orientation through direct pinch gestures. To encourage custom positioning, it is initially centrally positioned in front of them with a slight downward offset. Although repositioning affects self-visibility, it can be assumed that usually the user's head and shoulders are visible in the Personal Mirror. Due to its customization options while anchored to the local user, they can position it in a way to continuously observe themselves or only observe themselves on demand, for example, by looking down.

3.2.4 Miniature Avatar. This is a world-anchored self-view with high spatiality that represents the user as a scaled-down version of their 3D avatar, mirroring their body movements and facial expressions in real-time (see Figure 1D and 2D). It provides self-visibility of the user's upper body without restricting the perspective. Like the Personal Mirror, it can be scaled and repositioned in space using pinch gestures. However, after being placed, it remains stationary and does not follow the user. Initially, the Miniature Avatar is placed on the table in front of each participant, facing them to make them aware that it acts as a static object that does not follow the user. Depending on where the users place it, they can observe themselves almost continuously when placed in the FoV, only peripherally when placed further away from the FoV, or on demand when placed

Remote Perspective View: attached to partner's head | Remote Perspective View: second perspective | Personal Mirror: following the FoV | Personal Mirror: second perspective

Fig. 3. Screen captures from a first-person perspective of the Remote Perspective View (A&B), demonstrating how it is attached to the partner's head, and the Personal Mirror (C&D), demonstrating the attachment to the FoV – each by showing two perspectives on the same self-view.

out of view. Furthermore, the user can adjust the rotation of the Miniature Avatar to control its orientation. For example, the Miniature Avatar can face the user (i.e., viewing the front of the avatar), align with the user's perspective (i.e., viewing the back of the avatar), or be rotated differently (e.g., viewing the avatar from the side).

3.3 Avatar and Environment Design

For the avatar design, we aimed for a solution that offers customization options that represent a wide range of individuals, and is rather stylized to reduce potential uncanny valley effects [42, 53]. We opted for the Meta Avatars[7] (see Figure 1 and Figure 3) and used their web-based avatar editor[8] for customization. Participants were free to choose from all available customization options (e.g., facial features, body types, skin, hair, and eye colors) except for the outfit, which we set to a black shirt to increase comparability across sessions in terms of awareness and contrast to the background. In VR, participants embodied their avatars without legs because they are not tracked by the HMD, and the lack of movement (or only simulated movement) of the legs could potentially reduce embodiment. However, we compensated for this through our environment design, using a seated scenario, where participants were separated by a virtual table covering the non-existent lower halves of their bodies.

For the virtual room, we aimed for a neutral environment with some details in the background to make participants aware of how the flat self-views work. To meet these requirements, we adapted the 3D model of a living room[9] and added a table[10] in the middle (see Figure 4A). A physical table was placed in the real environment, and the virtual environment and table were calibrated to match it. During the study, both participants were in the same physical room and shared the same table. The participants' avatars were virtually placed at the same location where the participants were physically located. With this setup, we visually simulated a remote scenario in one physical room, circumventing the audio transmission implementation. This setup also enabled us to conduct the study in one lab without requiring multiple rooms and experimenters, which is common for most multi-user VR studies [64].

[7]Meta Avatars: https://meta.com/avatars/ (last accessed: May 30th, 2025)
[8]Meta Avatar Editor: https://quest.meta.com/avatar (account required, last accessed: May 30th, 2025)
[9]Living room with curtains: https://skfb.ly/o9roZ (last accessed: May 30th, 2025)
[10]Table: https://skfb.ly/pxFQZ (last accessed: May 30th, 2025)

3.4 Implementation

Our prototype was implemented using the Meta XR All-in-one SDK[11] in Unity3D[12] on the Meta Quest Pro[13] (106° × 95.57° FoV; 1800 × 1920 pixels per eye; hand, eye, and face tracking). Eye-tracking accuracy is reported as around 1.5° to 3° [5, 63]. For the integration of the Meta Avatars, we used the Meta Avatars SDK[14], which allows loading the avatar of the currently logged-in user into the scene, and apply the HMD's hand, eye, and face tracking data.

The self-views Remote Perspective View and Personal Mirror were implemented by rendering a mirrored camera image on a render texture. The Remote Perspective View has a size of 50 × 50 cm with a 2.5 cm white border around it. The virtual camera (60° FoV) capturing the image of the Remote Perspective View is placed at the head position of the remote avatar. However, the camera is not positioned at the center of the head but is offset by 50 cm toward the local user to make them appear larger in the view. The self-view screen is placed with an offset of 60 cm to the right and 10 cm up relative to the head of the remote avatar. In contrast, the Personal Mirror surface has an initial size of 20 × 20 cm with a 1.25 cm white border around it. The camera is placed behind the surface so that the FoV frustum (60°) matches the edges of the mirror surface to simulate the impression of real mirror proportions. The initial placement is 50 cm away and 20 cm below the local user's head. The Miniature Avatar was realized by additionally loading a scaled version of the own avatar into the scene. This avatar has an initial height of around 20 cm (depending on the selected hairstyle) and is initially placed on the table surface, looking at the local user. Both the Personal Mirror and Miniature Avatar can be scaled by a factor between 0.5 and 2, which corresponds to a size between 10 cm and 40 cm.

The coordinate systems of both devices in the room were synchronized using Meta's Shared Spatial Anchors[15]. The virtual table (and environment) was aligned with the physical table by calibrating the table surface with the HMD's controller. Communication between the devices was realized with Colibri [30] by exchanging control messages, avatar pose data, and additional alignment information.

4 User Study Comparing Self-Views

We conducted a user study to explore how users perceive having a self-view in VR. Based on related work, we defined three explorative research questions to gain a first understanding of how VR self-views are experienced.

RQ1: What are the user preferences and perceived usefulness of different self-view variants in VR?

Research shows that there is a preference for equal-sized video tiles [7] when using video conference tools and that seeing our virtual appearance affects our behavior in VR (e.g., [69]). Therefore, as we explore a new way of self-viewing in VR, we are interested in the participants' preferences and the perceived usefulness of the proposed self-views. To assess these aspects, we asked whether the self-view was distracting or helpful (based on the User Experience Questionnaire (UEQ) [36, 51]) at four defined points during each condition using a visual analog scale (VAS) [35] in VR. Additionally,

[11] Meta XR All-in-one SDK: https://developer.oculus.com/downloads/package/meta-xr-sdk-all-in-one-upm/ (last accessed: May 30th, 2025)
[12] Unity: https://unity.com/ (last accessed: May 30th, 2025)
[13] Meta Quest Pro: https://en.wikipedia.org/wiki/Meta_Quest_Pro (last accessed: May 30th, 2025)
[14] Meta Avatars SDK: https://developer.oculus.com/documentation/unity/meta-avatars-overview/ (last accessed: May 30th, 2025)
[15] Meta Shared Spatial Anchors: https://developer.oculus.com/documentation/unity/unity-shared-spatial-anchors/ (last accessed: May 30th, 2025)

in a concluding interview, we asked participants to rank all four conditions and which of the four options they would prefer for communication-focused or gesture-focused settings.

RQ2: How does the presence of a self-view in VR influence the self-perception & mental load of the user?

Self-views in video conferencing can foster self-consciousness, self-criticism, higher cognitive load [6], or social anxiety [41], while avatar self-representation in VR can enhance our perception of social situations [49] and confidence, e.g., during negotiations [69]. Thus, we are interested in whether a self-view in VR also affects the user's mental state. We asses mental load based on the respective question from the NASA Task Load Index [28] and the potentially mediating factor of task difficulty during this condition using a VAS in VR during four defined points in each condition.

Self-consciousness influences video conference situations. High self-consciousness is associated with social anxiety and low confidence about external perception [19]. Thus, as a ground truth, we initially assess public self-consciousness via the relevant subscale of the revised Self-Consciousness Scale (SCS-R) [50]. Since self-assurance/self-confidence has been linked to reducing social anxiety [3] and to counteracting self-consciousness [26], we also assess the participants' feelings of self-assurance on a scale from "very insecure" to "very confident" on a 7-point Likert scale.

Lastly, we assess presence using the SUS presence questionnaire [60] to control for any potential effects of the self-views on presence.

RQ3: How does the presence of a self-view in VR influence attention allocation between the self/self-view and interaction partners?

While we exhibit more natural communication behavior in VR [2], adding a self-view could disturb this by leading to an attention divide between the interaction partner and the self-view. Additionally, it was shown that depending on the task, we pay more or less attention to our mirrored image in VR [39]. Therefore, investigating self-views' influence on attention allocation is relevant in this context. We assess this by subjective, self-reported measurements on a questionnaire after each condition, utilizing three self-designed questions asking where participants allocated their attention during the last round (self-view, partner, and environment) and objectively by evaluating the participants' gaze throughout the study conditions and phases of the tasks.

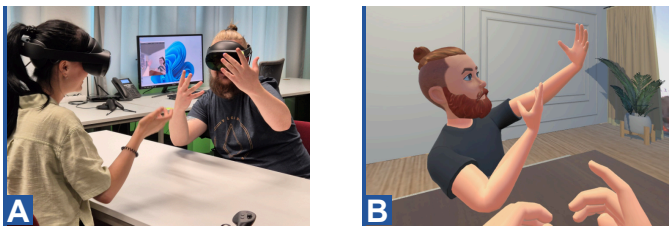
4.1 Task: Word-Guessing Game

We aimed to select an activity that covers a wide range of conversational scenarios, ranging from pure verbal communication to non-verbal communication using body gestures or facial expressions. Additionally, the task should be repeatable multiple times without getting boring or too exhausting. One possible scenario is to instruct participants to discuss different predetermined (polarizing) topics or to agree on a decision (as done for example by [29, 34, 46, 52, 56, 59]). However, we wanted to reduce emotional involvement and the complexity of the task to avoid influencing our measurements. One activity that meets our requirements is a word-guessing game, where a given term is explained either with words (similar to [15]) or by pantomiming it (similar to [4, 66]). This game has two easy rules: only use the predefined modality to explain the word (i.e., speaking or gesturing), and when speaking, do not say the word or parts of it. The game also allowed for rather casual interaction, and, because we only briefly showed the word in VR, the gameplay had a low risk of demanding all (visual) attention.

The participants played one round of the game for each self-view and one for the baseline. Each round consisted of four turns (2 minutes each), and each participant had a different role for each turn. First, Participant A had to explain a given word using only gestures and facial expressions (role: act-out), and Participant B had to guess the word (role: guess-during-act-out).

Virtual Environment | Setup Location One | Setup Location Two

Fig. 4. The virtual environment and the physical setup at the two different locations.



External View | Internal View

Fig. 5. Different views demonstrating the alignment of the real world with the virtual world.

After 2 minutes, Participant B had the act-out role and Participant A had to guess. After another 2 minutes, it was Participant A's turn again, who had to verbally explain a given word (role: explain), while Participant B had to guess (role: guess-during-explain). Finally, Participant B explained and Participant A guessed. Whenever a word was correctly guessed, a new one was given until the time limit of 2 minutes was reached. The order of words was fixed for all trials, but the lists were counterbalanced between all conditions to minimize their effect on the comparability of the experience. The English word lists can be found in the supplementary material.

4.2 Setup

The study was conducted in two labs of two different research facilities. As the study was done in immersive VR using two Meta Quest Pro HMDs, the virtual space at both labs was identical. The virtual space was a room containing a table, a window showing a static skyline, and some decorations such as plants (see Figure 4A, cf. section 3.3). The physical setup in both labs incorporated a table matching the width of the virtual counterpart (0.8 m × 1.2 m to 1.6 m, see Figure 5A and Figure 5B), where the participants sat on chairs facing each other at a distance of approx. 1 m (see Figure 4). The tables and avatars were aligned (cf. section 3.4). The remaining lab space varied (cf. Figure 4B and Figure 4C), but its influence on the task is assumed to be negligible. The setup further incorporated two PCs for the participants to create their avatars and fill out the questionnaires, as well as PCs for the instructor to monitor and control the study.

4.3 Procedure

We chose a within-subjects design with our three self-view variants (described in section 3.2) and the baseline as an independent variable. We used a balanced Latin square to counterbalance the

order of the four conditions: Remote Perspective View, Personal Mirror, Miniature Avatar, and No Add. Self-View (cf. section 3.2).

After giving informed consent, the participants filled out two questionnaires: (1) a demographic questionnaire to assess their self-reported gender, age, experience with video conference tools, preference for the self-view setting, experience with word-guessing games, and (2) the public self-consciousness subscale of the SCS-R. Next, the participants configured their avatar representations using the web-based avatar editor of Meta (cf. section 3.3) within a given time limit of 6 minutes. Following, the study instructor introduced the HMD and started the eye calibration procedure to ensure correct eye movements of the avatar and to accurately log the gaze. Before starting the tutorial phase of the session, the instructor needed to synchronize the coordinate systems of both participants and calibrate the table position. Then, in a short tutorial, the instructor explained the task procedure and where the words will appear in the virtual environment. After both participants indicated that they were familiar with the virtual environment and understood the procedure, they began with the actual tasks.

For each condition, participants first had a short period of 1-2 minutes to familiarize themselves with the self-view (or lack of an additional self-view). In the Miniature Avatar and Personal Mirror conditions, participants adjusted the self-view's position and scale during this phase. When finished, the task started, and participants played the word-guessing game. The four turns per condition started with Participant A acting out a given word and Participant B guessing, then taking turns, respectively (cf. section 4.1).

After each turn and before changing roles, both participants filled out an in-VR questionnaire about (1) the perceived difficulty of that turn, (2) the perceived mental workload, and (3) whether the self-view was distracting or helpful. This questionnaire was filled out by directly dragging the handle of a slider along a spectrum relating to the respective questions. When an answer was logged in, the next question would appear. When both participants were done, the next turn would start.

After completing a round that lasted around 15 minutes, the participants were instructed to remove their HMD, and take a break while filling out a post-condition questionnaire inquiring about different aspects of attention allocation and self-assurance, as well as the SUS presence questionnaire.

As soon as the participants felt ready to continue, they were instructed to return to the table and put on the VR HMD to repeat the same procedure with the next condition. In total, across four conditions, the participants each guessed eight times, acted out four times, and verbally explained the given words four times.

After completing all rounds, the post-study interview was conducted jointly with both participants. During the interview, we asked the participants about the likeliness of using each self-view in future VR sessions, asked them to rank all four conditions independently from one another on a sheet of paper, which of the four options they would prefer for communication-focused or gesture-focused settings, and general feedback. The questionnaires and the interview questions can be found in the supplementary material of this paper. In total, each study session lasted about 110 minutes, and participants were compensated adequately for their time. We followed all ethical and sanitary guidelines of our universities.

4.4 Participants

We recruited 16 participant dyads using non-probability sampling (8 male–male, 6 female–male, 1 female–female, and 1 non-binary–male). A prerequisite for participation was that each dyad already knew each other. This resulted in 32 participants (8 female, 23 male, and 1 non-binary) who were between 19 and 36 years old ($M = 23.59$, $SD = 3.89$). Eleven participants had corrected

vision with glasses, and one used contact lenses. Most participants ($n = 26$) had a background in computer science or closely related fields (computer engineering or media informatics), and the remaining ones had a background in psychology (3), mathematics (2), and speech and language processing (1). Most participants (20) did the study in their native language (German), while the others had sufficient proficiency to complete the study either in English (11) or the national language (German) (1), given the choice between both languages.

On a scale from 1 to 5, participants rated themselves as moderately experienced with AR ($M = 2.31, SD = 1.15$) and VR ($M = 2.38, SD = 1.31$), but only three used multi-user VR occasionally. They reported limited experience with Meta Avatars ($Median = 1, M = 1.53, SD = 0.80$). All participants were familiar with video conferencing tools, while most already used Zoom (32), video calls in Discord (30), WhatsApp video calls (28), and BigBlueButton (25). All but one reported having the self-view turned on during video calls. One participant had the self-view configured to be bigger than the others, 11 had it at the same size, and the majority (20) had it smaller than the other participants in the video conference. Out of 32 participants, only one said they do not play word guessing games, and three said they play it more than once a month.

5 Results

In the following, we report our thematically grouped results from the different measurements (questionnaires, gaze logging, and the final interview). As the gaze data is not normally distributed (assessed using QQ plots) and the questionnaire responses are on an ordinal scale, we analyzed our data using Friedman's ANOVA for repeated measures and Bonferroni-corrected Wilcoxon tests for post-hoc pairwise comparisons. We report Kendall's W for effect sizes.

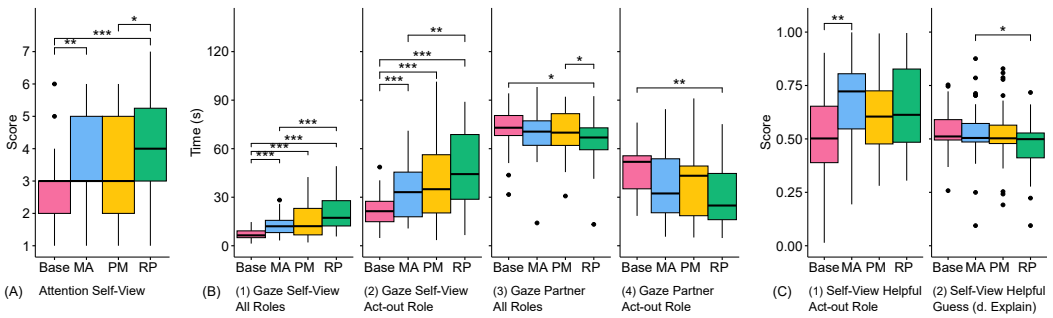


Fig. 6. Overview of study results. Abbreviations of the techniques in the figure: Base - No Add. Self-View, MA - Miniature Avatar, PM - Personal Mirror, RP - Remote Perspective View. **(A)** Post-condition questionnaire, significant results only. The plot shows the reported attention towards the self-views or own avatar on a scale of 1 (very low) to 7 (very high). **(B)** Comparison of the recorded gaze data. From left to right: 1. Time spent looking at the self-views or own avatar, aggregated over all four roles. 2. The same data just for the act-out role. 3. The time spent looking at the partner's avatar, aggregated over all roles. 4. The same data just for the act-out role. **(C)** Post-task questionnaire, significant results only. The plots show how helpful the self-view techniques were rated on a continuous scale from 0 (very distracting) to 1 (very helpful). Left: Ratings for the act-out role. Right: Rating for the guess-during-explain role. The boxes show the 25th and 75th percentiles, the horizontal bars mark the median, and the whiskers show the smallest and largest values up to 1.5 times the interquartile range; outliers are marked individually. Brackets mark significant differences (*: $p < .05$, **: $p < .01$, ***: $p < .001$)

5.1 Attention Allocation

To analyze attention allocation, we collected subjective questionnaire and objective gaze logging data. Responses to our post-condition questionnaires reveal highly significant differences in *reported attention to the self-views* ($\chi^2(3) = 19.73, p < .001$) between conditions (small effect size $W = 0.21$). Unsurprisingly, according to post-hoc tests users felt they paid significantly less attention to their embodied avatar in our baseline condition No Add. Self-View ($M = 2.63, SD = 1.21$) compared to the combination of embodied avatar and self-views in Miniature Avatar ($p < .01$, $M = 3.75, SD = 1.59$) and Remote Perspective View ($p < .001, M = 4.00, SD = 1.61$). Also, significantly lower attention scores were indicated for the Personal Mirror ($M = 3.22, SD = 1.79$) compared to Remote Perspective View ($p < .05$) (cf. Figure 6A).

To corroborate these findings, we also looked into the *recorded gaze data* of our participants (Figure 6B), whereby data of one participant had to be discarded due to tracking issues. While it is possible that these issues stem from the prescription glasses the participant wore during the study, we observed no such issues with any of the other ten participants with glasses. To this end, we compared the time spent looking directly at one's own embodied avatar alone (No Add. Self-View) to the combined time of looking at it directly or indirectly through the self-views (Miniature Avatar, Remote Perspective View, and Personal Mirror conditions).

Aggregated over all roles (described in section 4.1), we found a moderate effect of the variants on visual attention ($\chi^2(3) = 38.73, p < .001, W = 0.42$). Post-hoc tests showed that Remote Perspective View ($M = 19.67, SD = 10.48$) was viewed significantly longer ($p < .001$) on average compared to Miniature Avatar ($M = 12.67, SD = 6.71$). As visualized in Figure 6B_1, we also measured significantly less visual attention on the embodied avatar in our baseline condition No Add. Self-View ($M = 6.93, SD = 3.09$), compared to viewing it directly or through the self-view in Miniature Avatar ($p < .001$), Personal Mirror ($p < .001, M = 14.64, SD = 9.84$), and Remote Perspective View ($p < .001$) conditions. This partially confirms the tendencies in self-reported attention, with the Remote Perspective View being looked at most.

Examining the four individual user roles, we found that the time spent looking at one's own avatar in the baseline condition (No Add. Self-View) was consistently significantly lower than for the other conditions. This is expected, as the additional self-views do not replace the embodied avatar but supplement it. In three roles (act-out, guess-during-act-out, guess-during-explain), we found further significant differences between Miniature Avatar and Remote Perspective View that are in line with findings for the aggregated data. For example, this effect can be seen for the act-out role in Figure 6B_2, and further details can be found in out supplemental material.

Analyzing attention paid to the partner, we found no significant differences in *self-reported attention to the partner's avatar*, whereby the mean score for all four conditions is consistently high (No Add. Self-View: $M = 6.13, SD = 0.91$; Miniature Avatar: $M = 6.00, SD = 0.76$; Personal Mirror: $M = 6.06, SD = 0.88$; Remote Perspective View: $M = 6.03, SD = 0.60$). The *gaze data* shows a small but significant effect of the self-view variant on visual attention to the partner's avatar ($\chi^2(3) = 14.96$, $p < .01$, $W = 0.16$) with post-hoc tests (see Figure 6B_3) showing that participants looked less at their partner's avatar in the Remote Perspective View condition ($M = 65.30, SD = 14.83$) than in the No Add. Self-View baseline ($p < .05, M = 72.31, SD = 13.18$) and the Personal Mirror condition ($p < .05$, $M = 70.82, SD = 13.97$) . There was no significant difference to the Miniature Avatar condition ($M = 69.16, SD = 15.12$). Analyzing the roles individually, we only found a significant difference between conditions for the act-out role ($\chi^2(3) = 11.87$, $p < .01$, $W = 0.13$) , with post-hoc tests showing that participants looked less often at their partner in Remote Perspective View ($M = 31.17, SD = 19.52$) than in No Add. Self-View ($p < .01$,

$M = 46.21, SD = 15.10$) , as visualized in Figure 6B_4. This is expected when considering the time spent looking at the self-view in the same role, as depicted in Figure 6B_2.

5.2 Presence

We did not find a significant effect of the condition on the feeling of presence ($\chi^2(3) = 7.69, p > .05$), as reported in the post-condition questionnaire. Consequently, even though additional self-views may be seen as unrealistic (e.g., floating in space), there is no indication that our tested variants negatively influenced the sense of presence for our participants.

5.3 Mental Load, Task Difficulty and Helpfulness

We neither found a significant effect of the self-views on the mental load reported by our participants ($\chi^2(3) = 1.84, p > .05$) nor on how they perceived the difficulty of the words they had to explain or guess ($\chi^2(3) = 1.39, p > .05$). Asking whether the self-views were distracting or helpful during the task, we found no significant effect of condition for aggregated responses over all roles ($\chi^2(3) = 3.64$, $p > .05$). However, we found small but significant differences between conditions for the act-out role ($\chi^2(3) = 10.46, p < .05, W = 0.11$) and the guess-during-explain role ($\chi^2(3) = 13.88, p < .01$, $W = 0.14$) , as visualized in Figure 6C. Posthoc analysis reveals that participants assuming the act-out role found Miniature Avatar ($M = 0.68, SD = 0.19$) significantly more helpful than just the embodied avatar from the baseline (No Add. Self-View) ($p < .01$, $M = 0.51, SD = 0.22$), as can be seen in Figure 6C_1. Participants in the guess-during-explain role rated Miniature Avatar ($M = 0.53, SD = 0.15$) significantly better than Remote Perspective View ($p < .05$, $M = 0.47, SD = 0.13$) , as shown in Figure 6C_2. This data indicates that some self-view variants can be particularly helpful for gesture-focused tasks, but some designs may be more favorable.

5.4 Self-Assurance

Self-assurance reported in the post-condition questionnaire was similar across all conditions (Miniature Avatar: $M = 4.88, SD = 1.07$; Personal Mirror: $M = 4.38, SD = 1.24$; Remote Perspective View: $M = 4.63, SD = 1.19$; No Add. Self-View: $M = 4.53, SD = 1.05$) with no significant differences ($\chi^2(3) = 4.767, p > .05$) . The SCS-R results (SCS-R range [0;21]) range from 7 to 21 and are normally distributed (Shapiro-Wilk test $W = .982, p > .05$), with $M = 14.10$ and $SD = 3.04$. High values on the SCS-R indicate a high level of public self-consciousness, while high values for the self-designed post-condition question indicate a high level of self-assurance.

We analyzed the correlations between self-reported public self-consciousness (SCS-R) and self-reported self-assurance (post-condition questionnaire) for each variant. The results show a moderate (cf. [13]) positive correlation between the SCS-R and the self-assurance reports for Miniature Avatar (*Spearman's* $\rho = .330, p > .05$), a minor positive correlation for No Add. Self-View ($\rho = .103, p > .05$), a minor negative correlations for Remote Perspective View ($\rho = -.059$, $p > .05$) and Personal Mirror ($\rho = -.188, p > .05$), although none of the correlations are significant.

5.5 User Preferences and Feedback

In the post-study interview, participants were asked to rank each condition from most (1) to least preferred (4) (see Figure 7). Most preferred was the Miniature Avatar (average rank: 1.63, $SD = 0.79$), with 17 participants (53 %) ranking it as most preferred and an additional 11 participants (34 %) placing it as second best. The Personal Mirror was least preferred and placed last by 17 participants (53 %; average rank: 3.25; $SD = 0.95$). The other two techniques got an average rank of 2.5 (No Add. Self-View, $SD = 1.10$) and 2.63 (Remote Perspective View, $SD = 1.00$).

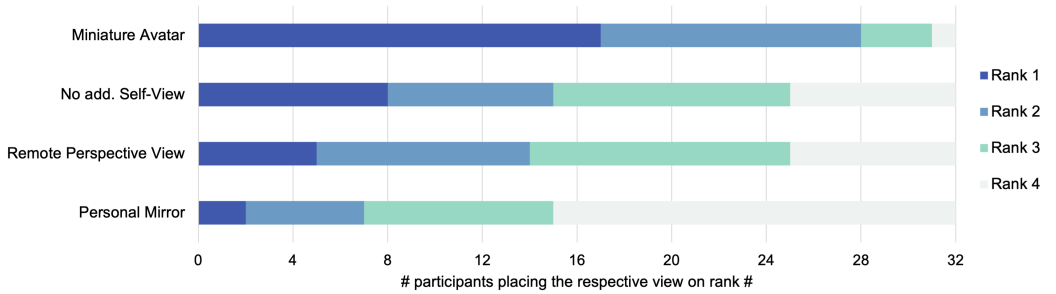


Fig. 7. The results from the participants' ranking of the four conditions (Miniature Avatar, Remote Perspective View, Personal Mirror, and No Add. Self-View) from most (rank 1) to least preferred (rank 4).

Miniature Avatar was the preferred self-view for most participants (23 for gestural tasks, 10 for conversational tasks), mostly due to the flexibility in positioning and visibility while interacting with their partner. This allowed participants to always see their complete representation without the need to move body parts into the view, compared to the other self-views. According to the participants, the 3D nature of the Miniature Avatar also supported them in checking their appearance and tracking, making it especially useful for gestural and physical tasks.

The Remote Perspective View has the advantage of being unobtrusive and consistent to the user while still allowing them to check the tracking, leading to five users preferring it for conversational tasks and four users for gestural tasks. However, not all participants perceived the positioning and perspective of this view to be optimal.

In contrast, the Personal Mirror was perceived as obtrusive and annoying, mainly because of its coupling to the users' body. Thus, only three participants would use it for gestural tasks, and none would prefer it for conversational tasks. Overall, participants wished for more customization options across all self-views, including the possibility of turning off the self-view when it is unnecessary (explicitly mentioned by 4 of 32).

These findings show that participants preferred the existence of additional self-views, as these allowed them to check the accuracy of the tracking and the appearance of oneself. When asked about preferences specifically for gestural tasks, No Add. Self-View was the least preferred option. However, appropriate design of the self-view is critical, as Personal Mirror was the least preferred option overall. Participants also reported not needing a self-view if they knew the tracking worked perfectly. Further, when considering conversational tasks specifically, 17 participants reported not needing a self-view, as these focus less on the partner's appearance.

Besides the specific feedback regarding the self-view variants, the participants reported feeling mostly represented by and liked the Meta avatars to different degrees. Only a minority (3 of 32) reported that drawbacks, like the unrealistic style, led to a negative opinion of the avatars. When asked why they looked at their self-representation, 20 participants reported wanting to check whether the body tracking worked correctly, and ten said they wanted to see explicitly how their gestures and facial expressions were shown to the other participant. Six of these participants had reported both. Twenty-one (21) participants also reported not noticing when their partners looked at their respective self-representation.

The participants were also asked if they would like to turn off certain movements or expressions (e.g., mouth movements, hand gestures) to others in some situations. Thirteen (13) participants would not want this at all, with some of them reasoning that this would be manipulative. Nineteen (19) participants could imagine it depending on the context and scenario. They claimed that

some information was unnecessary for a particular scenario and thus could be replaced with idle animations, or some movement could be suppressed to hide negative behaviors like nervous fidgeting or not paying attention to their conversation partner.

6 Discussion

In the following sections, we discuss our results and their limitations and provide recommendations for designing self-views in multi-user VR scenarios.

6.1 Helpfulness and User Preferences of Self-View Properties

Out of 32 participants, 24 (75 %) generally preferred to have any self-view over not having a self-view. Overall, 17 out of these 24 (ca. 71 %) prefer Miniature Avatar over all the other variants, and even out of the 8 people who would rather have no self-view, 5 (ca. 63 %) rated Miniature Avatar as their second favorite option. 30 out of 32 (ca. 94 %) participants noted they would want to have a self-view specifically during body-focused tasks, and out of those 30, 23 (ca. 77 %) found Miniature Avatar to be most suitable for these tasks.

Accordingly, our findings reveal that the Miniature Avatar was found to be significantly more helpful than No Add. Self-View during body-focused tasks. Participants highlighted the fact that Miniature Avatar, compared to Personal Mirror or Remote Perspective View, provided a better and more complete 3D impression of the tracked movement as a helpful characteristic, emphasizing the importance of providing high *Spatiality*, as described in 3.1. Participants expressed that they mostly looked at their self-view to monitor their gestures and facial expressions, verify if they appear as intended, experience what they look like from the other person's perspective, or verify if a specific movement was tracked and conveyed correctly. Participants liked that Miniature Avatar enables them to see their arms all the time, while for the other variants, hands and arms are usually out of frame unless lifted above chest level, pointing towards a preference for high *Self-Visibility*. Only 15 would prefer to have a self-view during solely conversational tasks, and out of those, 10 (ca. 67 %) named Miniature Avatar as their self-view of choice, and 5 (ca. 33 %) preferred Remote Perspective View. Here, both *Spatiality* and *Self-Visibility* seem to become less important.

Based on our data, we can extract the following preferences regarding *Anchoring and Size* of self-views: The self-views that were not head- or body-referenced to the user (Miniature Avatar, Remote Perspective View) were generally rated more favorably. Another frequently mentioned reason why Miniature Avatar was favored was its stationary, world-referenced position, that could be chosen freely. Participants often chose to place it unobtrusively outside of their immediate line of sight and, therefore, spent the least amount of time gazing at it, compared to the other self-views. This implies that opting out of looking at yourself is a favorable feature for VR self-views. Four participants explicitly said that they wished to be able to toggle the visibility of their self-views. The Remote Perspective View was the second most favored self-view, which was often explained by its *Anchoring and Size*, with its object-referenced position (here, the opposite user) being noted to be convenient, non-obstructive, and easily visible during the conversation. This comfortable yet prominent placement probably led to it being gazed at the most and the users spending less time looking at their partner's avatar compared to the other variants. Out of the three variants, the Personal Mirror was rated the least favorably, primarily due to its user-coupled placement, which users reported to intrude into their FoV. Prior work shows that users often prefer world-anchored over head-coupled content placed directly within their focus area [8], even when head-coupled solutions improve efficiency by shortening task completion times [24]. To reduce intrusiveness, we coupled this self-view only to the yaw-axis head rotation and allowed users to freely place it, even outside the immediate FoV, e.g., down low for on-demand viewing. Despite this flexibility, few

users leveraged the space above or below the focus area and instead placed it laterally at center height (see Figure 3C and Figure 3D), which still felt intrusive. Prior work also suggests that even peripherally placed virtual content near the body strongly attracts attention [55], making this self-view potentially more distracting than the other variants. Contrarily, in their self-reported attention distribution, users noted that they gave the Personal Mirror significantly less attention than the Remote Perspective View, however, this claim can not be verified when evaluating the gaze data, where no significant difference can be found. It can be assumed that Personal Mirror was often gazed upon unintentionally and subconsciously. However, according to the gaze data, it did not distract users significantly more or less than other self-views.

The participants were also asked if they had feedback or ideas for the current or future self-views. The responses were plentiful and varied greatly, indicating a need for individual customization options. Six participants would change the Personal Mirror to be world-anchored, three were fine with body-anchored, but wanted to place it further away than arm's length. Another wished for a different body-bound position, i.e., placing the personal mirror at their wrist, like a watch. Four wished to be able to dynamically adapt their self-view position throughout the conversation. Four participants wanted the Miniature Avatar to be full size, and place it either next to their conversation partner or replace them with it. Two wished for more self-visibility in the Remote Perspective View, two wished for the Miniature Avatar to have legs, and one said they would like to only see a face. Four wished to dynamically adapt the visibility and transparency of their self-views. Other mentioned adaptations were providing a pedestal for the Miniature Avatar to stand on, different visualization options (e.g., point clouds, silhouettes, doodles), providing a delayed mirror mode to allow watching the last couple of seconds of one's actions, and more.

In summary of the usefulness of self-views and the aforementioned user preferences (cf. RQ1), we propose that implementing self-views in multi-user VR should be strongly considered for gesture-focused tasks. Additionally, when designing self-views in VR, considerable efforts should be put towards evaluating their *Spatiality*, *Self-Visibility*, and *Anchoring and Size*. We conclude that self-views for gesture-focused tasks should be highly spatial, with sufficiently high self-visibility. They should be either world-anchored or object-anchored and either placed in a way that is not disruptive or distracting (while also being easily visible), or dynamically placeable by users themselves. We also recommend an opt-out function that enables users to control the visibility of their self-view, and suggest enabling further customization options for their position, scale, and visualization.

6.2 Influence of Self-Views on Self-Perception and Mental Load

We investigated how self-views in VR affect the mental state of the user by assessing their influence on self-perception and mental load (RQ2). Our results on mental load show no significant difference between conditions with additional self-views and the baseline condition with No Add. Self-View. Since task difficulty did not differ significantly between item sets – a potential mediating factor that could influence cognitive load – our results suggest that additional self-views did not increase mental load. Also, regarding the self-perception during the experience, participants consistently reported medium-high feelings of self-assurance (>4 on a 7-point Likert scale), with no significant differences between conditions, including No Add. Self-View. However, since self-assurance is affected by individual levels of public self-consciousness, we assessed participants' public self-consciousness in the beginning as a baseline. The results indicate that, overall, our sample was self-conscious about public perception, with some participants reporting high levels of self-consciousness. Since related work suggests that self-consciousness is usually negatively correlated with feelings of self-assurance [26], we expected a negative correlation between public self-consciousness and self-assurance, showing that high public self-consciousness leads to lower reported self-assurance. We found no significant correlation between self-consciousness and self-assurance for the individual

conditions. The found (expected) negative correlations were also only minor ($rho < .3$, cf. [13]) for the self-views Personal Mirror and Remote Perspective View. However, in contrast, for the Miniature Avatar, we found a moderate positive correlation, which could suggest that using the Miniature Avatar helped individuals with high public self-consciousness scores feel more confident. This is surprising and could suggest that Miniature Avatar may even have positively affected self-assurance, which can counteract self-consciousness [26] and thereby might help individuals feel less self-conscious and more confident in virtual meeting settings. Since this effect was not significant, future research should examine whether this trend can also be confirmed in different contexts. Generally, our results should be treated with some caution, as our measures represent only a subset of the reported negative aspects of self-monitoring during video conferencing [6], and future work needs to investigate in more detail whether the effects observed in video conferencing also apply to VR in general and in relation to self-views.

Based on our results, we conclude that adding self-views in VR does not affect the user's mental state in the same way as reported for video conferencing in related work [6, 41]. We found no negative effects for the specific aspects (i.e., mental load, self-assurance, and self-consciousness) we studied. A possible explanation could be that even though the avatar is personalized, it does not show the actual appearance of the person as in a video stream and gives the user control over their virtual appearance. On top of being independent of the user's actual visual appearance, the avatar representation is stable, i.e., its appearance does not change over time (e.g., no greasy hair or skin) and is not affected by the environment (e.g., hairstyle is not affected by wind), which is also one benefit of using avatars in meetings mentioned in related work [10]. Nevertheless, some participants explicitly wanted the ability to turn the self-view on and off as needed, suggesting that a permanent self-view is not always appreciated.

In summary, since we conclude that self-views in VR do not negatively impact mental load and self-perception and that Miniature Avatar even potentially reduces the burden of feeling self-conscious, we recommend adding a self-view in multi-user VR even for self-conscious individuals, but with the possibility to toggle its visibility.

6.3 Self-Views Influence (Perceived) Attention

As expected, we found an influence of additional self-views in VR on attention allocation during the tasks. The objective gaze data show that participants spent significantly more time looking at their own embodied avatar through additional self-views than directly at the avatar with No Add. Self-View. Similarly, with No Add. Self-View, on average, participants looked longer at their partner's avatar than when a self-view was visible (cf. section 6.1). In particular, with the Remote Perspective View, participants spent significantly less time looking at their partner's avatar and significantly more time focusing on the self-view, especially during the act-out role. Interestingly, when comparing all conditions with additional self-views, participants also spent significantly more time focusing on the Remote Perspective View than on the Miniature Avatar. This is most likely due to their respective positioning, since, as mentioned in 6.1, participants often placed their Miniature Avatar slightly out of view. We additionally asked the participants to self-report where they focused their attention. While the subjective results are mostly consistent with the objective gaze data, participants did not report paying significantly more attention to Personal Mirror compared to No Add. Self-View. Furthermore, participants subjectively reported a consistently high focus on the partner in all conditions ($Median = 6$ on a 7-point Likert scale), while objective gaze data show a small but significantly higher focus on the partner's avatar in the No Add. Self-View and Personal Mirror conditions compared to Remote Perspective View. This result might be affected by the position of Remote Perspective View, making it easy to glance at it

briefly while still feeling focused on the partner. When asked whether they noticed when their partner observed themselves in their self-view, 22 out of 32 denied it.

Overall, our results show that additional self-views significantly affect attention allocation between the partner's avatar and one's own embodied avatar perceived through the self-view. Among the additional self-view conditions, Remote Perspective View had the biggest impact on attention, while Miniature Avatar showed the smallest deviation from the baseline with No Add. Self-View. We recommend that a potential negative influence of diverted attention needs to be considered carefully, depending on the task. In critical cases, where users' visual attention must be mainly directed towards each other, self-views should be world-anchored and not intrude into the direct field of vision.

6.4 Design Recommendations

Based on the previous discussion, we provide design recommendations for self-views in VR while considering the limitations of our study, including the seated setting and the focus on only dual-user communication.

Provide Self-Views in VR for Gesture-focused Communication Scenarios. Since participants generally preferred self-views for gesture-focused tasks (cf. 6.1), we recommend integrating self-views into VR communication scenarios that require a high level of nonverbal communication using body movements such as gestures or facial expressions. Examples are casual communication with family, friends, or significant others (e.g., involving hand gestures such as waving, laughing, or discussing topics with a spatial nature, such as furniture arrangements) or guided practice and observational learning with an instructor (e.g., sign language).

Design Self-Views to Support Full Self-Visibility. Participants preferred the Miniature Avatar as it allowed them to view the entire body without considering the camera frame, as needed using our other self-views (cf. 6.1) or during video conferencing with a real camera (cf. [6]). Therefore, we recommend self-views in VR that provide *full self-visibility* of the user, e.g., by a representation with *high spatiality* as in the Miniature Avatar. However, we expect that variants with *low spatiality* such as Remote Perspective View will be able to provide this as well, e.g., by adaptively adjusting the displayed section in the frame based on user behavior, so hands or other essential body parts in those moments can always be observable without extra effort (cf. 6.1).

Give Users Control and Freedom to Customize Self-Views. Participants favored options to control the position and size across all self-views. Therefore, we recommend giving users the freedom to customize their self-view according to their personal preferences. In addition to position and size, options could also include enabling or disabling mirroring of the self-view (similar to video conferencing applications) or customizing the displayed section (i.e., self-visibility) in variants with *low spatiality*.

Carefully Design Self-View Positioning and Anchoring. In our seated scenario, participants were not comfortable with the Personal Mirror following the local user's perspective (cf. 6.1). Therefore, at least in seated scenarios, we recommend anchoring the self-view in the environment (world-anchored) with the freedom to change its position (as in the Miniature Avatar) or anchoring the self-view to the remote user representation (object-anchored, as in the Remote Perspective View). However, in scenarios where users are not bound to one position but move around the scene, we can only speculate that a self-view that dynamically follows the user might be beneficial but needs to be carefully designed.

Allow Use of Self-Views On Demand. Since some participants have requested the ability to turn self-views on and off as needed (cf. 6.1), we recommend always providing this option. However, we suggest that self-views are enabled by default in scenarios where they are beneficial (e.g., gesture-focused scenarios).

6.5 Limitations

While our study provides valuable insights into the usefulness and acceptability of self-views in VR, we acknowledge that it has certain limitations that may impact the generalizability and interpretation of our results. Our study involved only dyads, where the attention was only split between the self-view and the single conversation partner, while in group scenarios, the attention will be most likely split between multiple conversation partners. Therefore, we cannot directly scale our findings about the attention on the self-views to larger groups, but we expect a similar pattern, as the participants looked longer at their partner than at their self-view. Additionally, how the design of the opposite-user-anchored Remote Perspective View can be applied in group scenarios is not yet defined. One possibility would be to leverage the gaze and only show the Remote Perspective View next to the person one is currently looking at. As the study was conducted in a seating configuration to simulate meetings similar to ones done via video conferencing, the findings do not necessarily apply to a standing or mobile scenario. This especially applies to anchoring the self-view at a fixed position, as it might need to follow the user to a certain degree to be still observable. Additionally, our participants were co-located in the same physical room. This reduced the technical complexity of the setup and prevented issues with, e.g., audio streaming. On the other hand, this setup could influence communication behavior and affect the sense of presence as it provides natural spatial audio cues that would need to be simulated in a real remote scenario. Additionally, technical factors such as latency and tracking misalignments could have influenced the results, although the used server implementation has low latency and the setup was calibrated before each session. Our sample was not balanced in self-reported gender as it comprised only 8 female and 1 non-binary participants, while 23 identified as male. As previous research indicated that fatiguing effects might be stronger in female-identifying participants [18], it is important to seek a more balanced sample in future studies investigating self-views more deeply. While we found no indicators of fatigue based on our measurements, the task duration and task type could have been one factor in reducing the likeliness of its occurrence, as we had frequent breaks (roughly every 10 minutes) from VR and from the task itself. Additionally, the task frequently shifted the attention between listening, observing, and explaining in a predefined order, excluding some factors contributing to fatigue in video conference settings [1]. Furthermore, our chosen task relied on particular language knowledge to be able to describe the given words, either verbally or using gestures (pantomiming). Participants could choose to complete the study in either German or English. While most participants did the study in their native language (German), a subset did the study in English, which was not their native language in most cases. We can not rule out the possibility that this might have affected the perceived mental load. However, we attempted to measure this effect by asking participants to rate the perceived difficulty of the given words they had to explain or guess – an assessment likely influenced by their language proficiency – and found no significant differences across the self-view conditions.

7 Conclusion and Future Work

This work explores the effects of avatar self-views in multi-user VR scenarios, motivated by self-views in common video conferencing applications. Specifically, we designed and developed multiple self-view designs covering a range of design possibilities, choosing the following three based on pilot testing: (1) Remote Perspective View, (2) Personal Mirror, and (3) Miniature Avatar. We

compared them to a baseline condition with No Additional Self-View. Overall, our findings reveal that self-views are beneficial for communication scenarios that require expressive body language through gestures or facial expressions. Most participants preferred the Miniature Avatar for its flexibility in positioning and the overview of all avatar movements, while the Personal Mirror was the least preferred self-view and was perceived as obtrusive and annoying due to its anchoring to the local user. The Remote Perspective View was described as convenient and non-obtrusive but also distracted from the conversation partner the most. We condensed our findings into design recommendations for self-views in multi-user VR scenarios.

While our work is a first exploration revealing the benefits of self-views in VR, it is limited to a seated scenario and communication in dyads. Therefore, future work should further explore the influence of self-views in groups and scenarios involving movement in virtual space. Based on our user preference findings, further features of self-views can be explored in the future, such as enabling users to selectively toggle the visibility of individual body movements in VR, i.e., deactivating only gaze, gestures, or mouth movements during social interactions. Such customization options would greatly extend the common functionality known from video conferencing self-views, where usually only the video or audio transmission can be toggled. The design space around VR avatar self-views has yet to be exhaustively explored. With our initial exploration into the effect of self-views in multi-user VR, we lay the groundwork for future investigations of this previously under-explored area. Our findings and design recommendations suggest possible research directions, and with our work, we hope to inspire and inform the design of future user-centric social VR spaces.

Acknowledgments

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy: EXC 2117 – 422037984 (Centre for the Advanced Study of Collective Behaviour (CASCB)), as well as EXC 2050/1 – Project ID 390696704 – Cluster of Excellence "Centre for Tactile Internet with Human-in-the-Loop" (CeTI) of Technische Universität Dresden, and by DFG grant 389792660 as part of TRR 248 – CPEC (see https://cpec.science). The authors acknowledge the financial support by the Federal Ministry of Research, Technology and Space of Germany and by Sächsische Staatsministerium für Wissenschaft, Kultur und Tourismus in the programme Center of Excellence for AI-research „Center for Scalable Data Analytics and Artificial Intelligence Dresden/Leipzig", project identification number: ScaDS.AI. The authors further acknowledge the financial support by the Federal Ministry of Research, Technology and Space of Germany in the programme of "Souverän. Digital. Vernetzt.". Joint project 6G-life, project identification number: 16KISK001K. This work was also supported by the Alexander von Humboldt Foundation, funded by the German Federal Ministry of Research, Technology and Space. Lastly, the authors acknowledge the financial support of the Hector Foundation II. We thank Marc Satkowski for his insights and support and Uta Wagner for the inspiring discussions and valuable assistance with data analysis. We further want to thank Paul Rinau for his support in conducting the study.

References

[1] Jesper Aagaard. 2022. On the dynamics of Zoom fatigue. *Convergence* 28, 6 (2022), 1878–1891. doi:10.1177/13548565221099711

[2] Ahsan Abdullah, Jan Kolkmeier, Vivian Lo, and Michael Neff. 2021. Videoconference and Embodied VR: Communication Patterns Across Task and Medium. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW2 (Oct. 2021), 453:1–453:29. doi:10.1145/3479597

[3] Aneesh V Appu and Ms Vineetha Lukose. 2022. Objectified Body Consciousness and Social Interaction Anxiety: Self Confidence as a Moderator. *Indian Journal of Health Studies* 4 (2022), 07–21. doi:10.56490/IJHS.2022.4101

[4] Sahar Aseeri and Victoria Interrante. 2021. The Influence of Avatar Representation on Interpersonal Communication in Virtual Social Environments. *IEEE TVCG* 27, 5 (May 2021), 2608–2617. doi:10.1109/TVCG.2021.3067783

[5] Samantha Aziz and Oleg Komogortsev. 2022. An Assessment of the Eye Tracking Signal Quality Captured in the HoloLens 2. In *2022 Symposium on Eye Tracking Research and Applications (ETRA '22)*. ACM, New York, NY, USA, Article 5, 6 pages. doi:10.1145/3517031.3529626

[6] Jeremy N. Bailenson. 2021. Nonverbal Overload: A Theoretical Argument for the Causes of Zoom Fatigue. *Technology, Mind, and Behavior* 2, 1 (2021). doi:10.1037/tmb0000030

[7] Karolina Balogova and Duncan Brumby. 2022. How Do You Zoom?: A Survey Study of How Users Configure Video-Conference Tools for Online Meetings. In *Proc. of the 1st Annual Meeting of the Symposium on Human-Computer Interaction for Work (CHIWORK '22)*. ACM, New York, NY, USA, Article 5, 7 pages. doi:10.1145/3533406.3533408

[8] Manshul Belani, Harsh Vardhan Singh, Aman Parnami, and Pushpendra Singh. 2023. Investigating Spatial Representation of Learning Content in Virtual Reality Learning Environments. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 33–43. doi:10.1109/VR55154.2023.00019

[9] Mark Billinghurst, Hirokazu Kato, and Ivan Poupyrev. 2001. MagicBook: transitioning between reality and virtuality. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems (CHI EA '01)*. ACM, New York, NY, USA, 25–26. doi:10.1145/634067.634087

[10] Michael Bonfert, Anke V. Reinschluessel, Susanne Putze, Yenchin Lai, Dmitry Alexandrovsky, Rainer Malaka, and Tanja Döring. 2023. Seeing the Faces Is so Important—Experiences from Online Team Meetings on Commercial Virtual Reality Platforms. *Frontiers in Virtual Reality* 3 (Jan. 2023). doi:10.3389/frvir.2022.945791

[11] Mila Bujić, Anna-Leena Macey, Bojan Kerous, Anatolii Belousov, Oğuz Buruk, and Juho Hamari. 2023. Self-representation does (not) spark joy: Experiment on effects of avatar customisation and personality on emotions in VR. In *Proc. of the 7th International GamiFIN Conference (CEUR Workshop Proceedings)*. CEUR-WS, 96–105. https://ceur-ws.org/Vol-3405/paper9.pdf

[12] Ruizhi Cheng, Nan Wu, Songqing Chen, and Bo Han. 2022. Will Metaverse Be NextG Internet? Vision, Hype, and Reality. *IEEE Network* 36, 5 (Sept. 2022), 197–204. doi:10.1109/MNET.117.2200055

[13] Jacob Cohen. 2013. *Statistical Power Analysis for the Behavioral Sciences* (2 ed.). Routledge. doi:10.4324/9780203771587

[14] Tiffany D. Do, Camille Isabella Protko, and Ryan P. McMahan. 2024. Stepping into the Right Shoes: The Effects of User-Matched Avatar Ethnicity and Gender on Sense of Embodiment in Virtual Reality. *IEEE TVCG* 30, 5 (May 2024), 2434–2443. doi:10.1109/TVCG.2024.3372067

[15] Trevor J. Dodds, Betty J. Mohler, and Heinrich H. Bülthoff. 2011. Talk to the Virtual Hands: Self-Animated Avatars Improve Communication in Head-Mounted Display Virtual Environments. *PLOS ONE* 6, 10 (Oct. 2011), e25759. doi:10.1371/journal.pone.0025759

[16] Nina Döllinger, David Mal, Sebastian Keppler, Erik Wolf, Mario Botsch, Johann Habakuk Israel, Marc Erich Latoschik, and Carolin Wienrich. 2024. Virtual Body Swapping: A VR-Based Approach to Embodied Third-Person Self-Processing in Mind-Body Therapy. In *Proc. of the CHI Conference on Human Factors in Computing Systems (CHI '24)*. ACM, New York, NY, USA, 1–18. doi:10.1145/3613904.3642328

[17] Nina Döllinger, Erik Wolf, Mario Botsch, Marc Erich Latoschik, and Carolin Wienrich. 2023. Are Embodied Avatars Harmful to our Self-Experience? The Impact of Virtual Embodiment on Body Awareness. In *Proc of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, Hamburg Germany, 1–14. doi:10.1145/3544548.3580918

[18] G. Fauville, M. Luo, A. C. M. Queiroz, A. Lee, J. N. Bailenson, and J. Hancock. 2023. Video-conferencing usage dynamics and nonverbal mechanisms exacerbate Zoom Fatigue, particularly for women. *Computers in Human Behavior Reports* 10 (May 2023), 100271. doi:10.1016/j.chbr.2023.100271

[19] Allan Fenigstein, Michael F Scheier, and Arnold H Buss. 1975. Public and private self-consciousness: Assessment and theory. *Journal of consulting and clinical psychology* 43, 4 (1975), 522. doi:10.1037/h0076760

[20] Marie Luisa Fiedler, Erik Wolf, Nina Döllinger, Mario Botsch, Marc Erich Latoschik, and Carolin Wienrich. 2023. Embodiment and Personalization for Self-Identification with Virtual Humans. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 799–800. doi:10.1109/VRW58643.2023.00242

[21] Daniel Immanuel Fink, Moritz Skowronski, Johannes Zagermann, Anke Verena Reinschluessel, Harald Reiterer, and Tiare Feuchtner. 2024. There Is More to Avatars Than Visuals: Investigating Combinations of Visual and Auditory User Representations for Remote Collaboration in Augmented Reality. *Proc. ACM Hum.-Comput. Interact.* 8, ISS, Article 548 (Oct. 2024), 29 pages. doi:10.1145/3698148

[22] Maia Garau, Mel Slater, Simon Bee, and Martina Angela Sasse. 2001. The impact of eye gaze on communication using humanoid avatars. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '01)*. ACM, New York, NY, USA, 309–316. doi:10.1145/365024.365121

[23] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M. Angela Sasse. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 529–536. doi:10.1145/642611.642703

[24] Yalda Ghasemi, Ankit Singh, Myunghee Kim, Andrew Johnson, and Heejin Jeong. 2021. Effects of head-locked augmented reality on user's performance and perceived workload. In *Proceedings of the human factors and ergonomics society annual meeting*, Vol. 65. SAGE Publications Sage CA: Los Angeles, CA, 1094–1098. doi:10.1177/1071181321651169

[25] Mar González-Franco, Daniel Pérez-Marcos, Bernhard Spanlang, and Mel Slater. 2010. The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment. In *2010 IEEE Virtual Reality Conference (VR)*. IEEE, 111–114. doi:10.1109/VR.2010.5444805

[26] Stephen J Gould and Benny Barak. 1988. Public self-consciousness and consumption behavior. *The Journal of social psychology* 128, 3 (1988), 393–400. doi:10.1080/00224545.1988.9713756

[27] Jennifer A Harriger and Gabrielle N Pfund. 2022. Looking beyond zoom fatigue: The relationship between video chatting and appearance satisfaction in men and women. *International Journal of Eating Disorders* 55, 7 (2022), 923–932. doi:10.1002/eat.23722

[28] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (Oct. 2006), 904–908. doi:10.1177/154193120605000909

[29] Zhenyi He, Keru Wang, Brandon Yushan Feng, Ruofei Du, and Ken Perlin. 2021. GazeChat: Enhancing Virtual Conferences with Gaze-aware 3D Photos. In *The 34th Annual ACM Symposium on User Interface Software and Technology (UIST '21)*. ACM, New York, NY, USA, 769–782. doi:10.1145/3472749.3474785

[30] Sebastian Hubenschmid, Daniel Immanuel Fink, Johannes Zagermann, Jonathan Wieland, Harald Reiterer, and Tiare Feuchtner. 2023. Colibri: A Toolkit for Rapid Prototyping of Networking Across Realities. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 9–13. doi:10.1109/ISMAR-Adjunct60411.2023.00010

[31] Seoyoung Kang, Boram Yoon, Bowon Kim, and Woontack Woo. 2022. Effects of Avatar Face Level of Detail Control on Social Presence in Augmented Reality Remote Collaboration. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. IEEE, 763–767. doi:10.1109/ISMAR-Adjunct57072.2022.00161

[32] Do Yuon Kim, Ha Kyung Lee, and Kyunghwa Chung. 2023. Avatar-mediated experience in the metaverse: The impact of avatar realism on user-avatar relationship. *Journal of Retailing and Consumer Services* 73 (July 2023), 103382. doi:10.1016/j.jretconser.2023.103382

[33] Simon Kimmel, Frederike Jung, Andrii Matviienko, Wilko Heuten, and Susanne Boll. 2023. Let's Face It: Influence of Facial Expressions on Social Presence in Collaborative Virtual Reality. In *Proc. of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, New York, NY, USA, 1–16. doi:10.1145/3544548.3580707

[34] Jesper Kjeldskov, Jacob H. Smedegård, Thomas S. Nielsen, Mikael B. Skov, and Jeni Paay. 2014. EyeGaze: Enabling Eye Contact over Video. In *Proc. of the 2014 International Working Conference on Advanced Visual Interfaces (AVI '14)*. ACM, New York, NY, USA, 105–112. doi:10.1145/2598153.2598165

[35] Ludger Klimek, Karl-Christian Bergmann, Tilo Biedermann, Jean Bousquet, Peter Hellings, Kirsten Jung, Hans Merk, Heidi Olze, Wolfgang Schlenter, Philippe Stock, Johannes Ring, Martin Wagenmann, Wolfgang Wehrmann, Ralph Mösges, and Oliver Pfaar. 2017. Visual Analogue Scales (VAS): Measuring Instruments for the Documentation of Symptoms and Therapy Monitoring in Cases of Allergic Rhinitis in Everyday Health Care. *Allergo Journal International* 26, 1 (Feb. 2017), 16–24. doi:10.1007/s40629-016-0006-7

[36] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*, Andreas Holzinger (Ed.). Springer, Berlin, Heidelberg, 63–76. doi:10.1007/978-3-540-89350-9_6

[37] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. *3D user interfaces: theory and practice*. Addison-Wesley Professional.

[38] Irina Lediaeva and Joseph LaViola. 2020. Evaluation of Body-Referenced Graphical Menus in Virtual Environments. In *Graphics Interface 2020*. https://openreview.net/forum?id=k_qtvPF0QUJ

[39] Fang Ma and Xueni Pan. 2022. Visual Fidelity Effects on Expressive Self-avatar in Virtual Reality: First Impressions Matter. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 57–65. doi:10.1109/VR51125.2022.00023

[40] Yusuke Matsuda, Junya Nakamura, Tomohiro Amemiya, Yasushi Ikei, and Michiteru Kitazaki. 2021. Enhancing Virtual Walking Sensation Using Self-Avatar in First-Person Perspective and Foot Vibrations. *Frontiers in Virtual Reality* 2 (April 2021). doi:10.3389/frvir.2021.654088

[41] Matthew K. Miller, Martin Johannes Dechant, and Regan L. Mandryk. 2021. Meeting You, Seeing Me: The Role of Social Anxiety, Visual Feedback, and Interface Layout in a Get-to-Know-You Task via Video Chat.. In *Proc. of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, New York, NY, USA, 1–14. doi:10.1145/3411764.3445664

[42] Masahiro Mori, Karl F. MacDorman, and Norri Kageki. 2012. The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100. doi:10.1109/MRA.2012.2192811

[43] Margaret E Morris, Daniela K Rosner, Paula S Nurius, and Hadar M Dolev. 2023. "I Don't Want to Hide Behind an Avatar": Self-Representation in Social VR Among Women in Midlife. In *Proc. of the 2023 ACM Designing Interactive Systems Conference (DIS '23)*. ACM, New York, NY, USA, 537–546. doi:10.1145/3563657.3596129

[44] Niina Nurmi and Satu Pakarinen. 2023. Virtual meeting fatigue: Exploring the impact of virtual meetings on cognitive performance and active versus passive fatigue. *Journal of Occupational Health Psychology* 28, 6 (Dec. 2023), 343–362. doi:10.1037/ocp0000362

[45] Ken-Ichi Okada, Fumihiko Maeda, Yusuke Ichikawaa, and Yutaka Matsushita. 1994. Multiparty videoconferencing at virtual social distance: MAJIC design. In *Proc. of the 1994 ACM conference on Computer supported cooperative work (CSCW '94)*. ACM, New York, NY, USA, 385–393. doi:10.1145/192844.193054

[46] Kazuhiro Otsuka. 2018. Behavioral Analysis of Kinetic Telepresence for Small Symmetric Group-to-Group Meetings. *IEEE Trans. Multimed.* 20, 6 (June 2018), 1432–1447. doi:10.1109/TMM.2017.2771396

[47] Thammathip Piumsomboon, Gun A. Lee, Jonathon D. Hart, Barrett Ens, Robert W. Lindeman, Bruce H. Thomas, and Mark Billinghurst. 2018. Mini-Me: An Adaptive Avatar for Mixed Reality Remote Collaboration. In *Proc. of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 1–13. doi:10.1145/3173574.3173620

[48] Anca Salagean, Eleanor Crellin, Martin Parsons, Darren Cosker, and Danaë Stanton Fraser. 2023. Meeting Your Virtual Twin: Effects of Photorealism and Personalization on Embodiment, Self-Identification and Perception of Self-Avatars in Virtual Reality. In *Proc. of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, New York, NY, USA, 1–16. doi:10.1145/3544548.3581182

[49] Katsumi Sato and Yoko Usui. 2024. The Impact of Avatar Customization Flexibility on Con-versational Satisfaction. *IIAI Letters on Informatics and Interdisciplinary Research* 5 (Sept. 2024). doi:10.52731/liir.v005.242

[50] Michael F. Scheier and Charles S. Carver. 1985. The Self-Consciousness Scale: A Revised Version for Use with General Populations. *Journal of Applied Social Psychology* 15, 8 (1985), 687–699. doi:10.1111/j.1559-1816.1985.tb02268.x

[51] Martin Schrepp, Andreas Hinderks, and Jörg Thomaschewski. 2014. Applying the User Experience Questionnaire (UEQ) in Different Evaluation Scenarios. In *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience*, Aaron Marcus (Ed.). Springer International Publishing, Cham, 383–392. doi:10.1007/978-3-319-07668-3_37

[52] Martin Schuessler, Luca Hormann, Raimund Dachselt, Andrew Blake, and Carsten Rother. 2024. Gazing Heads: Investigating Gaze Perception in Video-Mediated Communication. *ACM Trans. Comput.-Hum. Interact.* 31, 3 (Aug. 2024), 39:1–39:31. doi:10.1145/3660343

[53] Valentin Schwind, Katrin Wolf, and Niels Henze. 2018. Avoiding the uncanny valley in virtual character design. *Interactions* 25, 5 (Aug. 2018), 45–49. doi:10.1145/3236673

[54] Sofia Seinfeld, Tiare Feuchtner, Antonella Maselli, and Jörg Müller. 2021. User representations in human-computer interaction. *Human–Computer Interaction* 36, 5-6 (2021), 400–438. doi:10.1080/07370024.2020.1724790

[55] Sofia Seinfeld, Tiare Feuchtner, Johannes Pinzek, and Jörg Müller. 2022. Impact of Information Placement and User Representations in VR on Performance and Embodiment. *IEEE Transactions on Visualization and Computer Graphics* 28, 3 (2022), 1545–1556. doi:10.1109/TVCG.2020.3021342

[56] Abigail J. Sellen. 1992. Speech Patterns in Video-Mediated Conversations. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '92)*. ACM, New York, NY, USA, 49–59. doi:10.1145/142750.142756

[57] Kristen M. Shockley, Allison S. Gabriel, Daron Robertson, Christopher C. Rosen, Nitya Chawla, Mahira L. Ganster, and Maira E. Ezerins. 2021. The fatiguing effects of camera use in virtual meetings: A within-person field experiment. *Journal of Applied Psychology* 106, 8 (Aug. 2021), 1137–1155. doi:10.1037/apl0000948

[58] Harrison Jesse Smith and Michael Neff. 2018. Communication Behavior in Embodied Virtual Reality. In *Proc. of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18)*. ACM, New York, NY, USA, 1–12. doi:10.1145/3173574.3173863

[59] Jian Sun and Holger Regenbrecht. 2007. Implementing Three-Party Desktop Videoconferencing. In *Proc. of the 19th Australasian Conference on Computer-Human Interaction: Entertaining User Interfaces (OZCHI '07)*. ACM, New York, NY, USA, 95–102. doi:10.1145/1324892.1324910

[60] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments* 9, 5 (Oct. 2000), 497–503. doi:10.1162/105474600566989

[61] Roel Vertegaal, Ivo Weevers, Changuk Sohn, and Chris Cheung. 2003. GAZE-2: conveying eye contact in group video conferencing using eye-controlled camera direction. In *Proc. of the SIGCHI Conference on Human Factors in Computing Systems (CHI '03)*. ACM, New York, NY, USA, 521–528. doi:10.1145/642611.642702

[62] Thomas Waltemate, Dominik Gall, Daniel Roth, Mario Botsch, and Marc Erich Latoschik. 2018. The Impact of Avatar Personalization and Immersion on Virtual Body Ownership, Presence, and Emotional Response. *IEEE TVCG* 24, 4 (April 2018), 1643–1652. doi:10.1109/TVCG.2018.2794629

[63] Shu Wei, Desmond Bloemers, and Aitor Rovira. 2023. A Preliminary Study of the Eye Tracker in the Meta Quest Pro. In *Proc. of the 2023 ACM International Conference on Interactive Media Experiences (IMX '23)*. ACM, New York, NY, USA, 216–221. doi:10.1145/3573381.3596467

[64] Xiaoying Wei, Xiaofu Jin, and Mingming Fan. 2024. Communication in Immersive Social Virtual Reality: A Systematic Review of 10 Years' Studies. In *Proc. of the Tenth International Symposium of Chinese CHI (Chinese CHI '22)*. ACM, New York, NY, USA, 27–37. doi:10.1145/3565698.3565767

[65] Erik Wolf, Marie Luisa Fiedler, Nina Döllinger, Carolin Wienrich, and Marc Erich Latoschik. 2022. 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)*. IEEE, 350–359. doi:10.1109/VR51125.2022.00054

[66] Yuanjie Wu, Yu Wang, Sungchul Jung, Simon Hoermann, and Robert W. Lindeman. 2021. Using a Fully Expressive Avatar to Collaborate in Virtual Reality: Evaluation of Task Performance, Presence, and Attraction. *Frontiers in Virtual Reality* 2 (April 2021). doi:10.3389/frvir.2021.641296

[67] Nick Yee and Jeremy Bailenson. 2007. The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research* 33, 3 (July 2007), 271–290. doi:10.1111/j.1468-2958.2007.00299.x

[68] Nick Yee and Jeremy N. Bailenson. 2009. The Difference Between Being and Seeing: The Relative Contribution of Self-Perception and Priming to Behavioral Changes via Digital Self-Representation. *Media Psychology* 12, 2 (May 2009), 195–209. doi:10.1080/15213260902849943

[69] Nick Yee, Jeremy N. Bailenson, and Nicolas Ducheneaut. 2009. The Proteus Effect: Implications of Transformed Digital Self-Representation on Online and Offline Behavior. *Communication Research* 36, 2 (April 2009), 285–312. doi:10.1177/0093650208330254

[70] Boram Yoon, Jae-eun Shin, Hyung-il Kim, Seo Young Oh, Dooyoung Kim, and Woontack Woo. 2023. Effects of Avatar Transparency on Social Presence in Task-Centric Mixed Reality Remote Collaboration. *IEEE TVCG* 29, 11 (Nov. 2023), 4578–4588. doi:10.1109/TVCG.2023.3320258

[71] Qiushi Zhou, Andrew Irlitti, Difeng Yu, Jorge Goncalves, and Eduardo Velloso. 2022. Movement Guidance using a Mixed Reality Mirror. In *Proceedings of the 2022 ACM Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 821–834. doi:10.1145/3532106.3533466

[72] Qiushi Zhou, Brandon Victor Syiem, Beier Li, Jorge Goncalves, and Eduardo Velloso. 2024. Reflected Reality: Augmented Reality through the Mirror. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 7, 4, Article 202 (Jan. 2024), 28 pages. doi:10.1145/3631431

Received 2025-03-14; accepted 2025-06-20