# Does Task Matter? Task-Dependent Effects of Cross-Device Collaboration on Social Presence

Christian Merz*
HCI & PIIS Group
University of Würzburg

Carolin Wienrich†
PIIS Group
University of Würzburg

Marc Erich Latoschik‡
HCI Group
University of Würzburg

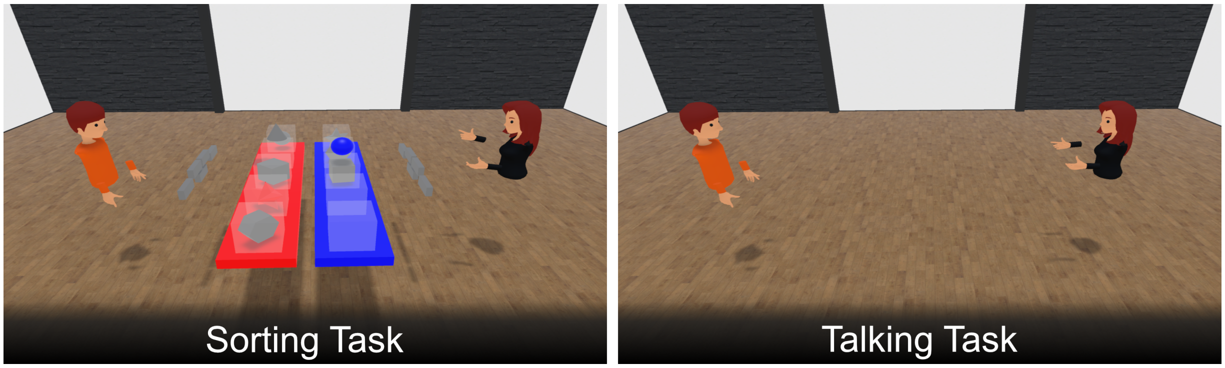

Figure 1: Third person perspective on the two participants during the respective tasks in the social XR environment with left *Sorting* and right *Talking*.

## ABSTRACT

In this work, we explored asymmetric collaboration under two distinct tasks: collaborative sorting and conversational talking tasks. We answer the research question of how different tasks impact the user experience in asymmetric interaction. Our mixed design compared one symmetric and one asymmetric interaction and two tasks, assessing self-perception (presence, embodiment), other-perception (co-presence, social presence, plausibility), and task perception (task load, enjoyment). 52 participants collaborated in dyads on the two tasks, either using head-mounted displays (HMDs) or one participant using an HMD and the other a desktop setup. Results indicate that differences in social presence diminished or disappeared during the purely conversational talking task in comparison to the sorting task. This indicates that differences in how we perceive a social interaction, which is caused by asymmetric interaction, only occur during specific use cases. These findings underscore the critical role of task characteristics in shaping users' social XR experiences and highlight that asymmetric collaboration can be effective across different use cases and is even on par with symmetric interaction during conversations.

**Index Terms:** Social presence, cross-device, asymmetric interaction, immersion, embodiment, social XR, social VR.

## 1 INTRODUCTION

Social virtual reality (VR) is a great tool for distributed collaboration and social interaction, offering the potential to replicate communication patterns as face-to-face collaboration through avatar representations [36]. It is a promising tool that can overcome the limitations of more established meeting platforms such as Microsoft Teams or Zoom that lack the ability to replicate the multidimensional aspects of in-person encounters, especially in conveying essential nonverbal, bodily, spatial, and, consequently, social cues.

However, challenges such as spatial constraints, health considerations, and hardware availability present significant barriers to widespread adoption that could prevent users from using it at all [38, 44]. Recent advances in commercial and research platforms aim to mitigate these barriers by enabling participation with various devices such as augmented reality (AR), desktop computers, or smartphones, effectively transforming the social VR platform into a social extended reality (XR) platform. This leads to asymmetric collaboration, in which users with different devices interact with each other [52]. Different devices lead to different user experiences [5]. Previous work has shown that device characteristics affect user experience, especially for spatial tasks in the 3D environment [2, 11, 45]. Nevertheless, how the different devices affect the complex dynamics of social interaction is an ongoing research topic, and underexplored [6, 52]. This is especially true when looking at balanced interaction design in asymmetric collaboration, where the interaction possibilities are not tailored to specific devices [22], such as when a desktop user acts as a guide for a VR user [10, 12]. A prior study has shown that replication of social cues - whether users can talk to each other - emphasizes the difference between the user experience with different devices. However, this was only evaluated for a single use case [23]. This demonstrates a lack of research on the balanced interaction in asymmetric collaboration. Therefore, in this work, we answer the following research question:

**RQ:** How do different tasks impact the user experience in asymmetric collaboration?

## 2 RELATED WORK

Immersion describes the technological capabilities that provide a simulation to a user's senses [35]. Because different devices (e.g., a head-mounted display and a desktop monitor) inherently offer different levels of immersion, users sharing the same virtual space may not have identical perceptual input. Therefore, they might experience the virtual environment differently [5].

Recent advances have facilitated multi-device participation in collaborative XR, allowing one user to join via a VR headset, while another joins via a desktop or even a mobile device [52, 22]. This

---

[*] e-mail: christian.merz@uni-wuerzburg.de
[†] e-mail: carolin.wienrich@uni-wuerzburg.de
[‡] e-mail: marc.latoschik@uni-wuerzburg.de

raises questions about how immersion differentials affect both one's own experience (e.g., sense of presence and embodiment) and one's impression of the interaction partner.

## 2.1 Asymmetric Interaction in Social XR

Research on cross-device applications often uses the term asymmetric interaction to indicate differences in users' roles or functionalities [6]. Prior work also refers to this as cross-reality when, for example, AR and VR users interact with each other [29, 37]. In this work, we refer to situations in which users participate under varying degrees of immersion while sharing the same virtual environment as asymmetric interaction [52].

Although previous work addresses asymmetric challenges by restricting certain actions for the less-immersed user (e.g., casting them as a "helper" or "observer" in VR) [10, 12, 26], many commercial platforms - such as RecRoom, Engage, or Meta Horizon Worlds - simply provide the same virtual space and interactions, regardless of the user's chosen device. In such cases, little is known about how immersion gaps affect collaboration quality and user experience when participants share the same roles and tasks [6, 43, 46]. Recent research addressed this by evaluating asymmetric interaction when users had the same interaction possibilities [22, 23]. These two studies evaluated the same task of sorting objects collaboratively, raising the question of how asymmetric interaction affects the user experience during different tasks. Previous work has shown that immersion can affect tasks like 3D docking or 3D visualization tasks [2, 11, 45]. However, there is a research gap that addresses how immersion affects simple conversations between users with different devices.

## 2.2 Effects of Immersion on User Experience

In evaluating the user experience within asymmetric social XR, several constructs help capture key indicators [17, 48].

However, self-perception (presence [33, 39, 41], embodiment [4, 27]), other-perception (co-presence, social presence [3, 53], plausibility [17, 20, 21, 34], and perceived humanness [9]), and task-perception (task load [8], task enjoyment and usability [19]) can be affected by the immersion of a device [22, 23].

### 2.2.1 Self-Perception

The spatial presence, which is described as the "sense of being there" in the virtual environment [35] - has been directly linked to immersion [5, 32, 35]. Devices offering larger fields of view, stereoscopic rendering, and higher tracking fidelity generally support a stronger sense of spatial presence [17, 35], showing that higher immersion leads to higher perceived spatial presence.

Users perceive a virtual avatar as their own body, which is called the sense of embodiment. The sense of embodiment incorporates ownership (feeling that a virtual body is "mine") and agency (feeling control over avatar actions) [14, 27]. The less immersive a setup is, the less users might experience visuomotor synchrony, potentially diminishing the perceived sense of embodiment [4, 7].

### 2.2.2 Other-Perception

Co-presence refers to the mutual awareness of other users and their actions [15], while social presence further addresses the psychological aspects of feeling connected to others in mediated spaces [3]. Previous work indicates that more immersive interfaces can enhance users' perceptions of "being together" [1, 30].

Another relevant construct is plausibility [17, 34]. The plausibility of a virtual human can be measured [20], which captures how convincing and natural avatars appear. Immersion affects how we perceive virtual humans and they may seem less plausible to the user [20, 21, 51], potentially affecting mutual social cues and engagement [16].

### 2.2.3 Task-Perception

Beyond socio-cognitive measures, task-related factors play a significant role in XR collaboration [8, 19]. More intuitive or higher-immersion interfaces can reduce effort and cognitive load, improving usability and enjoyment [40, 42]. In contrast, mismatches in immersion within the same environment might require additional effort for the lower-immersion user, potentially increasing perceived task load and lowering satisfaction [47].

In sum, there is currently a lack of research comparing different tasks in asymmetric interaction and how different tasks impact self-perception, other-perception, and task-perception. Therefore, we designed a study to evaluate the differences in these constructs when manipulating the asymmetry of the interaction and the specific use case.

# 3 METHOD

This study followed a 3 x 2 mixed model, where the between condition was the device configuration with either a symmetric or an asymmetric interaction which we refer to as the factor *Immersion*. The asymmetric interaction was a desktop user ($D_{VR}$) and a VR user ($VR_D$) interacting with each other. The symmetric interaction was two VR users ($VR_{VR}$) interacting with each other. For the naming of our device configurations, the first named device defines the device of the user, and the subscript defines the device of the interaction partner. The within condition was the factor *Task* the participants had to solve, with either the collaborative sorting task (*Sorting*) where the data is from a prior publication [23], and the other task was two users talking about various topics (*Talking*). The device configuration was balanced. However, the task order was always in the order of first *Sorting* and second *Talking*.

## 3.1 Procedure

We recruited two participants and placed each individual in a separate room for remote collaboration. A researcher escorted each participant to their assigned experimental space. As illustrated in Figure 2, the study began with participants signing a consent form and reviewing the study information. They then completed an initial set of questionnaires, including demographics and control measures, before entering the private virtual environment using their assigned devices. In this private space, participants were able to explore the interface and customize their avatars according to personal preferences. After finalizing their avatar designs, they transitioned to a shared social virtual environment to work together on a sorting task. Once the task was finished, they filled out post-experiment questionnaires. Afterwards, participants got again study information for the second task (*Talking*). Participants were able to customize their avatars again and had a short acclimatization phase in a private virtual environment. Then they again joined the social virtual environment where they were tasked to talk to their partner about suggested topics or topics of their own choice for ten minutes. Following the *Talking* task participants had to fill out the post-questionnaires again. To maintain consistency, we synchronized the timing of each step across both participants. Overall, the study took approximately 60 minutes to complete.

## 3.2 System Description

Our system was created using Unity3D (version 2020.3.21f1), with network operations managed by Photon's PUN2 toolkit. Rather than a peer-to-peer architecture, we implemented a client-to-server approach to streamline the distribution of information: the system collects data packets from each client at set intervals, processes them on the server, and then dispatches the updates to the other participants. By running the information exchange at a 20Hz frequency and incorporating lag-reduction mechanisms, the system can efficiently transmit each user's avatar state (e.g., position, orientation, and appearance) along with other data like task progress to every

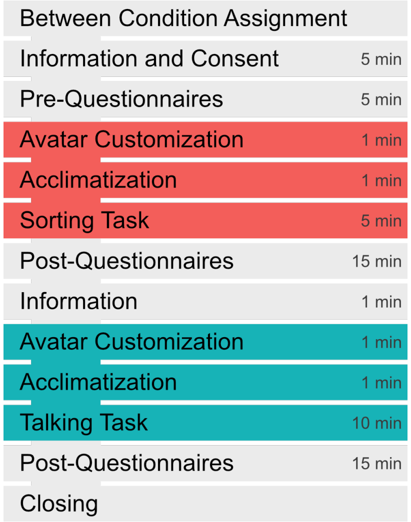

Figure 2: The figure shows the procedure of our study, where red highlights the parts in the virtual environment from our earlier publication of the *Sorting* task [23] and turquoise highlights the *Talking* task.

client. Prior to collaborating together, each participant could select a stylized male or female avatar and customize the appearance (e.g., skin, hair, and shirt).

In the virtual environment, our avatars are partially truncated between the hips and shoulders/hands—often referred to as a "rayman style"—so as to avoid the complexities of full-body inverse kinematics, particularly when only three tracking devices are available [49]. Figure 1 presents an illustration of how these avatars appear in our platform.

For the *VR* setup, we used the HTC Vive Pro Eye headset, which offers a 110-degree field of view, a 90Hz refresh rate, and a per-eye resolution of 1440×1600 pixels, paired with two Vive Pro Controllers. On the desktop side (*D*), participants interacted via a 27-inch monitor (1920×1080, 60Hz) and a standard office mouse. We integrated the SteamVR framework (version 1.14.15) into our Unity environment to bring in positional data from both the HMD and the controllers. Device positions were tracked by a pair of HTC Base Stations 2.0 using an infrared-based system, which sustains a 22ms latency and sub-millimeter accuracy [24] at a 1000Hz sampling rate. Each room had its own set of base stations mounted on tripods for stable and comprehensive coverage.

Across both conditions (*VR* and *D*), participants wore identical Lenovo 100 Stereo USB headsets to maintain consistent audio quality between devices. The study ran on powerful desktop machines with Windows 10, each featuring an Intel i7-11900K CPU, NVIDIA GeForce RTX 3080 GPU, and 64GB of DDR5 RAM, providing ample performance for the experimental requirements.

## 3.3 Task and System Design

Figure 3 illustrates the perspectives of both participants as they carried out the task in the *Sorting* and *Talking* task. If participants wished to change their viewpoint in the virtual environment, they had to hold down the secondary mouse button and drag the mouse to rotate their perspective in the *D* condition. Meanwhile, in the *VR* condition, participants used Vive controllers that tracked their hand movements, and they were able to look around naturally by moving their heads while wearing the VR headset.

### 3.3.1 Sorting Task

During the *Sorting* task each participant stood at a differently colored table—one blue and one red. Hovering above each table were five partially transparent containers, into which participants needed to sort a set of five objects in ascending order, determined by the number of corners on each object.

At the start of each round, a new object appeared above the container at the center of the table. Participants used a pair of buttons labeled "Left" and "Right" to shift the object's position, then used a "Confirm" button to place it into the chosen container and prompt the arrival of the next object. In cases where an object matched the partner's table color, participants transferred it to the other side by pressing an "Interact" button. Selecting "Interact" on one side caused the partner's corresponding button to change color, signaling that a transfer request was pending. Upon the partner's acceptance, the item would appear on their table. The process continued until all objects were appropriately sorted. When participants pressed the "Finish" button, they concluded the task. Buttons that were not currently available to use appeared in black.

We aimed to maintain a uniform interaction style across the different devices. Participants only interacted with the system through button presses. In the *D* setting, they used a mouse to click on the virtual buttons. Each click activated a pre-recorded animation of a virtual hand pressing the button. In the *VR* condition, participants had to press the buttons by moving their hands to the virtual buttons.

### 3.3.2 Talking Task

In the *Talking* task, participants stood in front of each other and were tasked to their interaction partner about topics of their preference or the following proposed topics: Favourite food, favorite movies, remote lectures, weekend activities, or their last vacation.

Participants could not move in the virtual environment but only look around. Participants in the *VR* condition could move their virtual hands since their real hands were tracked, while there was no tracking for the hands of the participants in the *D* condition. The task concluded when 10 minutes passed.

## 3.4 Measures

We used immersive tendency and simulator sickness as control variables, reflecting their importance in XR research [28]. We measured immersive tendency via the ITQ with 18 items on a 7-point scale [50], and simulator sickness via the SSQ with 14 items on a 0-4 scale [13].

To evaluate self-perception, we used the Igroup Presence Questionnaire (IPQ) [31], which contains 19 items spanning the subdimensions of spatial presence, involvement, and realism. We also used the Virtual Embodiment Questionnaire (VEQ) [27] to measure the sense of embodiment, encompassing ownership, agency, and change.

Other-perception was measured with the Networked Mind Measurement (NMM) [3], focusing on co-presence and social presence. NMM includes four subscales—two related to co-presence for oneself and one's partner, and two evaluating psychobehavioral engagement for oneself and one's partner. Additionally, the Virtual Human Perception (VHP) questionnaire [20] provided subscales examining appearance and behavioral plausibility (ABP) as well as alignment with the virtual environment (MVE). We further included one single-item measure of the perceived human likeness of the interaction partner, rated 1–7.

Task perception in this study was evaluated using the Raw NASA Task Load Index (RTLX) [8]. We also incorporated two single-item measures where participants provided 1–7 ratings of both their enjoyment of the task and the overall usability of the system.

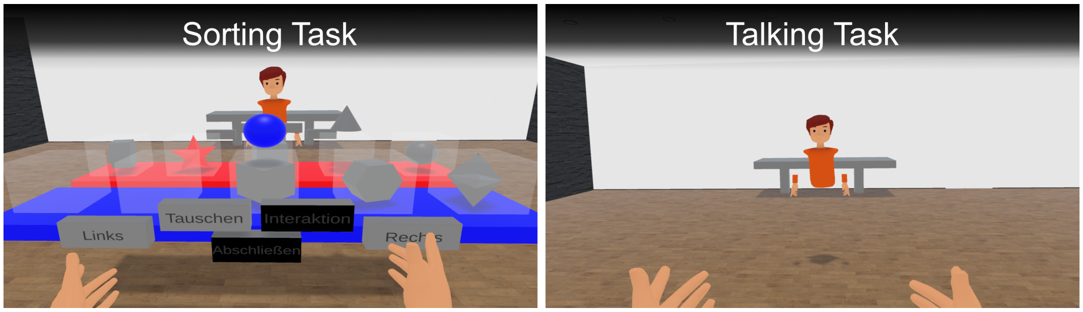

Figure 3: The figure shows the first person perspective during our study, for left the *Sorting* task and right the *Talking* task.

### 3.5 Participants

Our sample size included $N = 52$ participants who took part in a prior study [23]. $N = 49$ were students who received credit points as part of their bachelor's degree, and $N = 3$ received money as compensation. The students had the same knowledge background of studying media and computer science. We had to exclude two participants, as technical difficulties arose during the interaction in the virtual environment. The condition $D_{\text{VR}}$ and $VR_{\text{D}}$ had $N = 16$ participants, and $VR_{\text{VR}}$ had $N = 18$ participants. Our sample of $N = 50$ was $M = 20.36$ $(SD = 1.84)$ years old. 47 were female, 3 were male. 80% of the participants had at least one hour of VR experience.

## 4 RESULTS

We used Python 3.9 with the Numpy and Pandas libraries to aggregate data and compute scores for our results. For the statistical analysis, we used R 4.4.0 with RStudio 2024. Multiple measurements showed violations of the mixed ANOVA's normality and variance homogeneity assumptions. Since we had a relatively small sample size with balanced groups, we calculated a robust variation with the trimmed means ANOVA with a 20% cut-off based on the R package WRS2 [18]. To validate our results, we calculated a nonparametric ANOVA with the *f1.ld.f1* method of the *nparLD* R package [25] that showed similar results for the analysis. Therefore, we only report the results of the trimmed means ANOVA for our variables. In the upcoming sections, we will report only significant results for the main and interaction effects and significant post hoc tests when the main effects are significant. Additionally, we only report post hoc tests for the main effect of immersion and the interaction effect since the main effect of the task already is a comparison between two groups, and the post hoc test does not reveal further information. We applied the Bonferroni correction for the post hoc tests to keep the family-wise error small. Table 1 shows the descriptive values for all dependent variables.

### 4.1 Control Measures

Simulator sickness was not significantly different between conditions, $F(4) = 0.44, p = .779$. However, it differed significantly between the time of measurements - before the experiment (T1), after the first task (T2), and after the second task (T3) - $F(2) = 96.07, p < .001$. However, the score decreased over time, shown with the following means: T1, M = 14.38 (SD = 4.07); T2, M = 10.40 (SD = 2.81); T3, M = 9.18 (SD = 2.18).

The immersive tendency of the participants showed no significant differences between the devices, $F(2) = 0.26, p = .768$.

### 4.2 Self-Perception

#### 4.2.1 Presence

Figure 4 shows a bar plot with confidence intervals for the subscale of spatial presence.

Spatial Presence A two-way ANOVA revealed a significant main effect of *Immersion*, $F(2, 16.89) = 9.87, p = .002$, and a significant main effect of *Task*, $F(1, 21.74) = 9.25, p = .006$, without interaction effect ($p = .438$). Post hoc comparisons indicated that participants reported higher spatial presence under $VR_{VR}$, $t(94) = -5.49, p < .001$, than under $D_{VR}$. Spatial presence was significantly higher for *Sorting* than *Talking*.

Involvement There was a significant main effect of *Immersion*, $F(2, 17.05) = 17.27, p < .001$, while the effect of *Task* and the interaction remained non-significant. In post hoc tests, involvement was higher for $VR_{VR}$ than for $D_{VR}$, $t(94) = -4.88, p < .001$.

Realism There were no significant main or interaction effects.

#### 4.2.2 Sense of Embodiment

Figure 5 and Figure 6 show bar plots with confidence intervals for perceived ownership and agency.

Ownership The analysis showed significant main effects of *Immersion*, $F(2, 17.80) = 4.38, p = .028$, and *Task*, $F(1, 20.37) = 5.89, p = .025$. Post hoc comparisons for *Immersion* revealed higher ownership under $VR_{VR}$ than $D_{VR}$, $t(94) = -3.51, p = .002$. Ownership was significantly higher for *Sorting* than *Talking*. The interaction was not significant.

Agency There were significant main effects of *Immersion*, $F(2, 16.12) = 23.60, p < .001$, and *Task*, $F(1, 16.97) = 4.83, p = .042$. Post hoc tests for *Immersion* indicated that $VR_{VR}$ yielded higher agency than $D_{VR}$, $t(94) = -7.71, p < .001$. Perceived agency was significantly higher for *Sorting* than *Talking*. The interaction was not significant.

Change There were no significant main or interaction effects.

### 4.3 Other-Perception

#### 4.3.1 Co-Presence

There was no significant effect for either dimension of co-presence.

#### 4.3.2 Social Presence

Figure 7 and Figure 8 show bar plots with confidence intervals for the two scales of social presence.

Psychobehavioral Engagement Self There was a main effect of *Task*, $F(1, 22.84) = 28.80, p < .001$. The main effect of *Immersion* was not significant. There was a tendency for the interaction effect to differ between conditions, $F(1, 21.64) = 31.49, p < .001$. Therefore, we calculated the post-hoc tests anyway, even though the p-value was not lower than our threshold of .05. Post hoc analysis of the interaction indicated that *Talking* induced higher other-engagement than *Sorting* for $D_{VR}$, $t(94) = -2.84, p = .039$, and $VR_D$, $t(94) = -3.33, p = .009$. During *Sorting*, participants with $VR_{VR}$ had higher scores than participants using $D_{VR}$, $t(94) = -3.10, p = .018$.

| | | | $D_{\text{VR}}$ | | $VR_{\text{D}}$ | | $VR_{\text{VR}}$ | |
|---|---|---|---|---|---|---|---|---|
| | | | *Sorting* | *Talking* | *Sorting* | *Talking* | *Sorting* | *Talking* |
| Measure | Subscale | Sign. | N=16 | N=16 | N=16 | N=16 | N=18 | N=18 |
| IPQ | Spatial Presence | *I* & *T* | 3.88 (1.13) | 3.55 (1.22) | 5.32 (1.00) | 4.89 (1.13) | 5.40 (0.65) | 4.69 (0.79) |
| | Involvement | *I* | 3.56 (1.31) | 3.09 (1.30) | 5.11 (1.10) | 5.09 (1.15) | 5.08 (1.51) | 4.62 (1.28) |
| | Realism | | 3.73 (0.54) | 3.83 (0.69) | 3.91 (0.77) | 3.88 (0.66) | 3.69 (0.53) | 3.81 (0.72) |
| VEQ | Ownership | *I* & *T* | 2.16 (1.22) | 1.75 (0.90) | 3.5 (1.41) | 3.00 (1.80) | 3.38 (1.51) | 2.93 (1.41) |
| | Agency | *I* & *T* | 3.58 (1.41) | 2.73 (1.46) | 5.52 (1.06) | 5.34 (1.00) | 5.74 (1.02) | 5.28 (1.47) |
| | Change | | 2.26 (2.06) | 1.44 (1.19) | 3.03 (1.54) | 2.35 (1.34) | 2.60 (1.83) | 2.28 (1.10) |
| NMM | CP self | | 3.78 (0.68) | 4.03 (1.01) | 3.91 (0.71) | 4.19 (0.64) | 3.94 (0.48) | 3.81 (0.53) |
| | CP other | | 3.83 (0.75) | 3.43 (0.74) | 3.84 (0.38) | 3.84 (0.36) | 3.88 (0.42) | 3.72 (0.41) |
| | PE self | *T* | 3.30 (0.86) | 4.17 (1.12) | 3.64 (0.82) | 4.67 (0.60) | 4.23 (0.87) | 4.61 (0.88) |
| | PE other | *I*T* | 3.08 (1.16) | 4.39 (0.98) | 3.71 (1.03) | 4.56 (0.69) | 4.41 (0.96) | 4.57 (0.80) |
| VHP | ABP | | 5.05 (0.92) | 5.12 (1.22) | 5.04 (0.95) | 5.30 (1.04) | 5.55 (0.68) | 5.62 (0.75) |
| | MVE | | 5.75 (0.97) | 5.72 (1.02) | 5.84 (1.02) | 5.89 (0.84) | 6.18 (0.72) | 6.24 (0.67) |
| Item | Humaness | *I* | 4.06 (2.32) | 5.12 (1.86) | 4.62 (1.54) | 5.44 (1.41) | 5.17 (1.65) | 5.11 (1.53) |
| RTLX | Mental dem. | *T* | 42.50 (22.86) | 9.38 (13.65) | 43.75 (23.13) | 20.31 (22.25) | 46.39 (18.61) | 10.83 (11.28) |
| | Physical dem. | *T* | 16.56 (16.10) | 3.12 (5.44) | 20.62 (16.42) | 5.94 (10.36) | 16.94 (12.85) | 1.39 (2.87) |
| | Frustration | *T* | 23.12 (12.76) | 5.62 (11.95) | 41.88 (30.60) | 16.88 (27.92) | 25.56 (16.79) | 8.33 (19.40) |
| | Temp. dem. | *T* | 22.50 (17.89) | 17.81 (19.66) | 31.25 (26.86) | 15.00 (27.81) | 29.72 (21.11) | 9.44 (16.71) |
| | Performance | *T* | 39.06 (27.16) | 18.44 (19.81) | 44.38 (28.74) | 24.38 (30.82) | 35.83 (24.99) | 11.67 (16.89) |
| | Effort | *T* | 26.25 (17.18) | 6.25 (8.27) | 27.81 (25.43) | 11.88 (10.63) | 34.44 (24.61) | 8.61 (13.04) |
| Item | Usability | *T* | 4.56 (1.55) | 6.12 (0.75) | 3.75 (1.81) | 6.12 (1.09) | 5.00 (1.19) | 6.22 (0.65) |
| | Task Joy | *I* & *T* | 5.06 (1.12) | 6.25 (0.86) | 5.81 (1.05) | 5.88 (1.36) | 6.00 (0.77) | 6.56 (0.86) |

Table 1: This table shows the descriptive data with mean, and standard deviation for the three different device conditions with the *Sorting* task of Merz et al. [23] and the *Talking* task. An *I* indicates a significant main effect for *immersion,* and a *T* indicates a significant main effect for *task.* If there is a significant interaction effect, we indicate it with a *I*T*, and we do not highlight the significant main effects.

**Psychobehavioral Engagement Other** There were significant main effects of *Immersion*, $F(2,17.96) = 4.52, p = .026$, and *Task*, $F(1,21.64) = 31.49, p < .001$, accompanied by a significant *Immersion* × *Task* interaction, $F(2,17.33) = 9.60, p = .002$. Post hoc analysis of the interaction indicated that *Talking* induced higher psychobehavioral engagement other than *Sorting* for $D_{VR}$, $t(94) = -3.93, p = .001$, and $VR_D$, $t(94) = -2.53, p = .049$. During *Sorting*, $VR_{VR}$ had significantly higher values than $D_{VR}$, $t(94) = -4.09, p < .001$.

### 4.3.3 Virtual Human Plausibility

The two subscales of VHP - appearance behavior plausibility and match to the virtual environment - showed no significant effect.

### 4.3.4 Humanness

There was a main effect of *Task* on humanness, $F(1,21.03) = 4.39, p = .049$, showing that the perceived humanness was significantly higher in the *Talking* than in the *Sorting* task. However, descriptive values show a similar effect as the social presence subscales indicating that a higher power with a larger sample size could reveal an interaction effect.

## 4.4 Task Perception

**Usability.** There was a significant main effect of *Task*, $F(1,19.75) = 40.87, p < .001$, participants rated usability higher in the *Talking* task than in the *Sorting* task. *Immersion* and the interaction did not differ significantly.

**Task Enjoyment.** Analyses revealed a main effect of *Immersion*, $F(2,16.74) = 4.55, p = .027$, and a main effect of *Task*, $F(1,16.57) = 19.59, p < .001$, with no significant interaction. *Talking* showed significantly higher enjoyment than *Sorting*. Post hoc comparisons showed that $VR_{VR}$ led to higher enjoyment than $D_{VR}$, $t(94) = -2.52, p = .040$.

### 4.4.1 Task Load

**Mental Demand** There was a significant main effect of *Task*, $F(1,26.87) = 70.61, p < .001$, showing that *Sorting* was more mentally demanding than *Talking*. *Immersion* and the interaction were not significant.

**Physical Demand** *Task* had a significant effect, $F(1,25.65) = 35.57, p < .001$, indicating that *Sorting* was rated higher in physical demand than *Talking*. No other effects were significant.

**Frustration** A main effect of *Task* was found, $F(1,13.58) = 47.16, p < .001$, revealing greater frustration during *Sorting* than *Talking*, $t(94) = 4.73, p < .001$. The *Immersion* factor and the interaction were not significant.

**Temporal Demand** *Task* again showed a significant effect, $F(1,22.24) = 26.28, p < .001$, where *Sorting* was more temporally demanding than *Talking*. There was no significant main effect on *Immersion* and no interaction effect.

**Performance** A main effect of *Task*, $F(1,22.64) = 30.44, p < .001$, indicated poorer perceived performance under *Sorting* than *Talking*. No other effects were significant.

**Effort** Finally, *Task* exhibited a main effect, $F(1,21.80) = 22.37, p < .001$, *Sorting* required significantly more effort than *Talking*. *Immersion* and the interaction showed no significance.

## 5 DISCUSSION

In this work, we present a 2 × 3 mixed study by manipulating the asymmetry and the task participants had to do in social XR. We measured key experience indicators that evaluate how the participants perceived themselves, the other, and the task.

In the second task (*Talking*), participants did not have to interact spatially in the virtual environment but had to simply talk to their

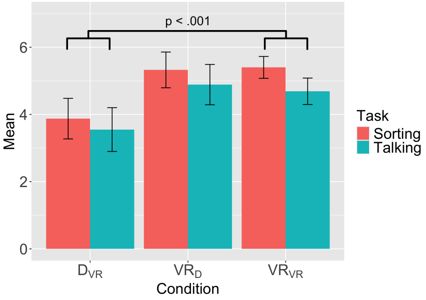

Figure 4: Mean values and CI(±95%) for perceived spatial presence with a significant main effect on *task* and *immersion*.

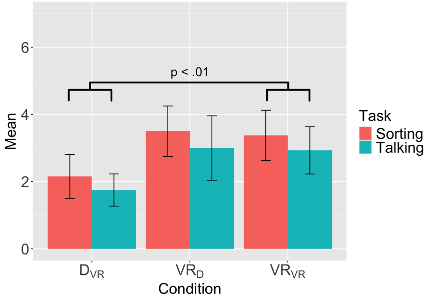

Figure 5: Mean values and CI(±95%) for ownership with a significant main effect on *task* and *immersion*.

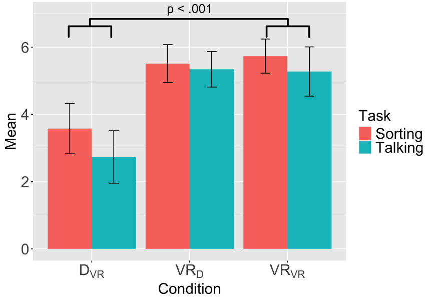

Figure 6: Mean values and CI(±95%) for perceived agency with a significant main effect on *task* and *immersion*.

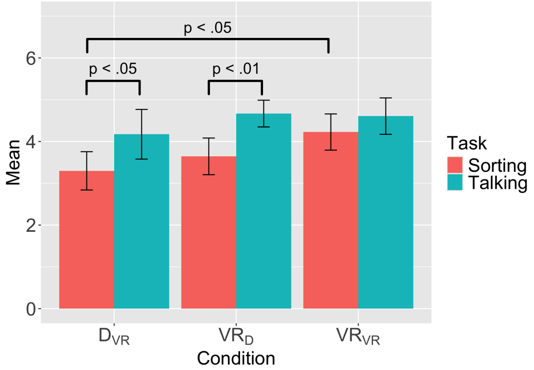

Figure 7: Mean values and CI(±95%) for perceived self psychological engagement.

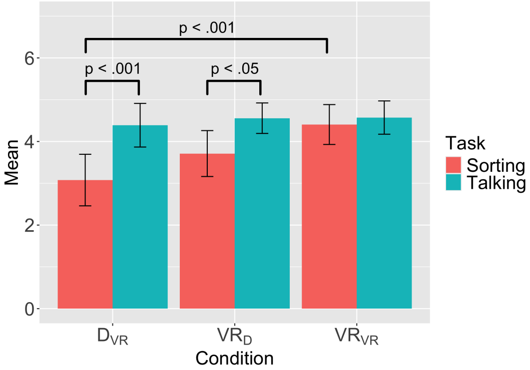

Figure 8: Mean values and CI(±95%) for other psychological engagement.

interaction partner. Hence, no interactions required specific movement where the users would see their virtual body but the nonverbal communication (body movement). However, the desktop users could not see their nonverbal movement - except head movement - since the hands were not tracked. The results reflect this since spatial presence and ownership were significantly lower for the *Talking* task compared to the *Sorting* task, where participants interacted spatially with their avatar representation. Additionally, spatial presence was higher for users in the *VR* condition than for users in the *D* condition. This is in line with previous work that immersion affects spatial presence [5, 35]. Therefore, we can assume that our manipulation worked as expected during both tasks. The results of ownership and agency support the findings and the argumentation for the results of spatial presence. Again, this shows that our manipulation worked as intended as visuomotor synchrony is higher for the *VR* condition, which is, in turn, a strong predictor for the sense of embodiment [14, 27]. Visuomotor synchrony might be higher in the *Sorting* task than in the *Talking* task since participants did not have to interact with their virtual hands during the *Talking* task. However, we can only assume this since we did not record the movement of the participants or the eye tracking to verify that in the *VR* condition, visuomotor synchrony was lower for the participants.

The interaction effect for the two subscales of social presence shows that after the second task, there are no significant differences between the device configurations, and we could no longer find the previous effect of the asymmetric device configurations. There are two different interpretations of the interaction effects of social presence. First, we could have found an order effect that simply by solving a task together and being together in social XR for a longer time, the participants reported overall higher scores for social presence in the *Talking* task. The implication would then be that the difference in other-perception and collaboration disappear simply by collaborating for a longer time. This means that the top-down effects of longer exposure or social interaction might overshadow the bottom-up effects of less immersion. However, we suspect that there should have been an increase in social presence for the VR$_{VR}$ condition as well. Therefore, we argue that the second interpretation is more valid and based on the two tasks' different affordances. In the *Talking* task, participants merely needed to converse rather than interact spatially within the virtual environment. As a result, they were not required to move around, and participants focused more on talking with their partners than recognizing the limitations of the desktop configuration with less immersion. This task characteristic likely led to the diminished effect in the social presence of the asymmetric device configurations in this specific scenario.

Therefore, in both interpretations, the asymmetry might impact the social interaction only in specific use cases or disappear after longer collaboration and interaction. Recent work showed that there are only differences in social presence when adding social cues [23]. Considering the results of this work, there seems to be a more complex dynamic when evaluating asymmetric interaction since social cues and task type can influence social presence. Therefore, when designing these systems, developers have to consider these two factors and how they might impact their application.

The use case did impact how the task load was perceived. Overall, the sorting task was rated as more demanding. However, the task load was not significantly different when using different devices. This indicates that asymmetry can facilitate effective collaboration across different use cases. This is in line with prior work that evaluated asymmetric interaction and demonstrated that even though there are differences in perception when using different device configurations, it is an effective form of social interaction and collaboration [10, 22, 23]

### 5.1 Limitations & Future Work

Although our results shed light on how different devices affect user experience across multiple use cases, our evaluation holds limitations. First, our sample size was relatively small and homogeneous, which restricts the generalizability of our findings. While the groups were balanced, larger-scale studies involving diverse participant demographics and contexts (e.g., corporate versus academic environments) are needed to confirm and extend our results. Second, the tasks we selected—sorting objects and engaging in casual conversation—were relatively simple and highly controlled. More complex or domain-specific tasks (e.g., collaborative design or data visualization) might introduce additional interaction requirements, potentially altering how asymmetry influences social presence and user experience in general. Especially when looking at other tasks that were evaluated when manipulating immersion that were more focused on spatial interaction in the virtual environment [2, 11, 45]. Moreover, our strict ordering (*Sorting* first, then *Talking*) could have introduced an order effect. Although our results suggest an interplay between use case and device configuration, it would be informative to counterbalance the task order in future studies to definitively disentangle the effect, whether it stems from longer exposure or the type of task. Future work could broaden the interaction repertoire to examine how more sophisticated or varied input modalities might affect social presence, task load, and collaborative efficiency. Despite these limitations, our study demonstrates the importance of investigating asymmetry more deeply, prompting new questions about how device-specific affordances and task complexity interplay to shape user experience and, specifically, social presence.

## 6 CONCLUSION

This study explored the influence of device asymmetry on user experience in two different use cases in social XR. We found that discrepancies in social presence diminished or disappeared entirely in a purely conversational context. This suggests that task type is a critical moderating factor in asymmetric collaboration. Furthermore, we found that device asymmetry did not significantly affect the perceived task load in either scenario, implying that collaboration remains feasible and effective even when participants operate different devices with varying levels of immersion. These findings underscore the potential of designing social XR platforms that accommodate a wide spectrum of hardware configurations without compromising user experience. Overall, our work advances the understanding of how different devices affect social presence, embodiment, and task perception within XR environments, especially when users engage in balanced or uniformly designed interactions. Future investigations may expand to more complex tasks and diverse input methods to develop comprehensive guidelines for designing effective and inclusive asymmetric XR systems.

## ACKNOWLEDGMENTS

We thank Tana Glemser for her help in conducting the study. This research has been funded by the German Federal Ministry of Education and Research (BMBF) in the project HiAvA (Grant 16SV8781) and by the Bavarian State Ministry for Digital Affairs in the project XR Hub (Grant A5-3822-2-16).

## REFERENCES

[1] D. Ahn, Y. Seo, M. Kim, J. H. Kwon, Y. Jung, J. Ahn, and D. Lee. The effects of actual human size display and stereoscopic presentation on users' sense of being together with and of psychological immersion in a virtual character. *Cyberpsychology, Behavior, and Social Networking*, 17(7):483–487, 2014. doi: 10.1089/cyber.2013.0455 2

[2] L. Besançon, P. Issartel, M. Ammi, and T. Isenberg. Mouse, tactile, and tangible input for 3d manipulation. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 4727–4740, 2017. doi: 10.1145/3025453.3025863 1, 2, 7

[3] F. Biocca, C. Harms, and J. Gregg. The networked minds measure of social presence: Pilot test of the factor structure and concurrent validity. In *4th annual international workshop on presence, Philadelphia, PA*, pp. 1–9, 2001. 2, 3

[4] F. Born, S. Abramowski, and M. Masuch. Exergaming in vr: The impact of immersive embodiment on motivation, performance, and perceived exertion. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pp. 1–8, 2019. doi: 10.1109/VS-Games.2019.8864579 2

[5] J. J. Cummings and J. N. Bailenson. How immersive is enough? a meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19(2):272–309, 2016. doi: 10.1080/15213269.2015.1015740 1, 2, 6

[6] B. Ens, J. Lanir, A. Tang, S. Bateman, G. Lee, T. Piumsomboon, and M. Billinghurst. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81–98, 2019. doi: 10.1016/j.ijhcs.2019.05.011 1, 2

[7] D. Gall, D. Roth, J.-P. Stauffert, J. Zarges, and M. E. Latoschik. Embodiment in virtual reality intensifies emotional responses to virtual stimuli. *Frontiers in Psychology*, 12, 2021. doi: 10.3389/fpsyg.2021.674179 2

[8] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909 2, 3

[9] C.-C. Ho and K. F. MacDorman. Measuring the uncanny valley effect: Refinements to indices for perceived humanness, attractiveness, and eeriness. *International Journal of Social Robotics*, 9:129–139, 2017. doi: 10.1007/s12369-016-0380-9 2

[10] R. Horst, R. Dörner, and M. Peter. Vr-guide: A specific user role for asymmetric virtual reality setups in distributed virtual reality applications. 10 2018. 1, 2, 7

[11] H. H. Huang, H. Pfister, and Y. Yang. Is embodied interaction beneficial? a study on navigating network visualizations. *Information Visualization*, 22(3):169–185, 2023. doi: 10.1177/14738716231157082 1, 2, 7

[12] K. Jeong, J. Kim, M. Kim, J. Lee, and C. Kim. Asymmetric interface: user interface of asymmetric virtual reality for new presence and experience. *Symmetry*, 12(1):53, 2020. doi: 10.3390/sym12010053 1, 2

[13] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The international journal of aviation psychology*, 3(3):203–220, 1993. doi: 10.1207/s15327108ijap0303_3 3

[14] K. Kilteni, R. Groten, and M. Slater. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387, 2012. doi: 10.1162/PRES_a_00124 2, 6

[15] L. Kohonen-Aho and P. Alin. Introducing a Video-Based Strategy for Theorizing Social Presence Emergence in 3D Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 24(2):113–131, 05 2015. doi: 10.1162/PRES_a_00222 2

[16] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology*, VRST '17. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3139131.3139156 2

[17] M. E. Latoschik and C. Wienrich. Congruence and plausibility, not presence: Pivotal conditions for xr experiences and effects, a novel approach. *Frontiers in Virtual Reality*, 3:694433, 2022. doi: 10.3389/frvir.2022.694433 2

[18] P. Mair and R. Wilcox. Robust statistical methods in r using the wrs2 package. *Behavior research methods*, 52:464–488, 2020. 4

[19] G. Makransky and G. B. Petersen. Investigating the process of learning with desktop virtual reality: A structural equation modeling approach. *Computers & Education*, 134:15–30, 2019. doi: 10.1016/j.

compedu.2019.02.002 2

[20] D. Mal, E. Wolf, N. Döllinger, M. Botsch, C. Wienrich, and M. E. Latoschik. Virtual human coherence and plausibility – towards a validated scale. In *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 788–789, 2022. doi: 10.1109/VRW55335.2022.00245 2, 3

[21] D. Mal, E. Wolf, N. Döllinger, M. Botsch, C. Wienrich, and M. E. Latoschik. From 2d-screens to vr: Exploring the effect of immersion on the plausibility of virtual humans. In *CHI 24 Conference on Human Factors in Computing Systems Extended Abstracts*, p. 8, 2024. doi: 10.1145/3613905.365077 2

[22] C. Merz, C. Göttfert, C. Wienrich, and M. E. Latoschik. Universal access for social xr across devices: The impact of immersion on the experience in asymmetric virtual collaboration. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 859–869. IEEE, 2024. doi: 10.1109/VR58804.2024.00105 1, 2, 7

[23] C. Merz, C. Wienrich, and M. E. Latoschik. Does voice matter? the effect of verbal communication and asymmetry on the experience of collaborative social xr. In *2024 IEEE international symposium on mixed and augmented reality (ISMAR)*. IEEE, 2024. 1, 2, 3, 4, 5, 6, 7

[24] D. C. Niehorster, L. Li, and M. Lappe. The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research. *i-Perception*, 8(3):2041669517708205, 2017. doi: 10.1177/2041669517708205 3

[25] K. Noguchi, Y. R. Gel, E. Brunner, and F. Konietschke. nparld: An r software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software*, 50(12):1–23, 2012. doi: 10.18637/jss.v050.i12 4

[26] T. Piumsomboon, G. A. Lee, J. D. Hart, B. Ens, R. W. Lindeman, B. H. Thomas, and M. Billinghurst. Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pp. 1–13, 2018. doi: 10.1145/3173574.3173620 2

[27] D. Roth and M. E. Latoschik. Construction of the virtual embodiment questionnaire (veq). *IEEE Transactions on Visualization & Computer Graphics*, 26(12):3546–3556, 2020. doi: 10.1109/TVCG.2020.3023603 2, 3, 6

[28] S. Rózsa, R. Hargitai, A. Láng, A. Osváth, E. Hupuczi, I. Tamás, and J. Kállai. Measuring immersion, involvement, and attention focusing tendencies in the mediated environment: The applicability of the immersive tendencies questionnaire. *Frontiers in Psychology*, 13:931955, 2022. doi: 10.3389/fpsyg.2022.931955 3

[29] J.-H. Schröder, D. Schacht, N. Peper, A. M. Hamurculu, and H.-C. Jetter. Collaborating across realities: Analytical lenses for understanding dyadic collaboration in transitional interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*, pp. 1–16, 2023. 2

[30] R. Schroeder. Being there together and the future of connected presence. *Presence*, 15(4):438–454, 2006. doi: 10.1162/pres.15.4.438 2

[31] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. doi: 10.1162/105474601300343603 3

[32] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2:647997, 2021. doi: 10.3389/frvir.2021.647997 2

[33] M. Slater. Measuring presence: A response to the witmer and singer presence questionnaire. *Presence*, 8(5):560–565, 1999. doi: 10.1162/105474699566477 2

[34] M. Slater, D. Banakou, A. Beacco, J. Gallego, F. Macia-Varela, and R. Oliva. A separate reality: An update on place illusion and plausibility in virtual reality. *Frontiers in Virtual Reality*, 3, 2022. doi: 10.3389/frvir.2022.914392 2

[35] M. Slater and S. Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 6(6):603–616, 1997. doi: 10.1162/pres.1997.6.6.603 1, 2, 6

[36] H. J. Smith and M. Neff. Communication behavior in embodied virtual reality. vol. 2018-April. Association for Computing Machinery, 4 2018. doi: 10.1145/3173574.3173863 1

[37] L. Stacchio, A. Angeli, and G. Marfia. Empowering digital twins with extended reality collaborations. *Virtual Reality & Intelligent Hardware*, 4(6):487–505, 2022. 2

[38] J.-P. Stauffert, F. Niebling, and M. E. Latoschik. Latency and cybersickness: impact, causes, and measures. a review. *Frontiers in Virtual Reality*, 1:582204, 2020. doi: 10.3389/frvir.2020.582204 1

[39] J. Takatalo, T. Kawai, J. Kaistinen, G. Nyman, and J. Häkkinen. User experience in 3d stereoscopic games. *Media Psychology*, 14(4):387–414, 2011. doi: 10.1080/15213269.2011.620538 2

[40] N. Talukdar and S. Yu. Breaking the psychological distance: the effect of immersive virtual reality on perceived novelty and user satisfaction. *Journal of Strategic Marketing*, pp. 1–25, 2021. doi: 10.1080/0965254X.2021.1967428 2

[41] R. Tamborini and P. Skalski. The role of presence in the experience of electronic games. In *Playing video games*, pp. 263–281. Routledge, 2012. 2

[42] Q. Tang, Y. Wang, H. Liu, Q. Liu, and S. Jiang. Experiencing an art education program through immersive virtual reality or ipad: Examining the mediating effects of sense of presence and extraneous cognitive load on enjoyment, attention, and retention. *Frontiers in Psychology*, 13, 2022. doi: 10.3389/fpsyg.2022.957037 2

[43] W. Tong, M. Xia, K. Wong, D. A. Bowman, T. Pong, H. Qu, and Y. Yang. Towards an understanding of distributed asymmetric collaborative visualization on problem-solving. In *2023 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*, pp. 387–397. IEEE Computer Society, Los Alamitos, CA, USA, mar 2023. doi: 10.1109/VR55154.2023.00054 2

[44] S. Vlahovic, M. Suznjevic, N. Pavlin-Bernardic, and L. Skorin-Kapov. The effect of vr gaming on discomfort, cybersickness, and reaction time. In *2021 13th International Conference on Quality of Multimedia Experience (QoMEX)*, pp. 163–168. IEEE, 2021. doi: 10.1109/QoMEX51781.2021.9465470 1

[45] X. Wang, L. Besançon, M. Ammi, and T. Isenberg. Understanding differences between combinations of 2d and 3d input and output devices for 3d data visualization. *International Journal of Human-Computer Studies*, 163:102820, 2022. doi: 10.1016/j.ijhcs.2022.102820 1, 2, 7

[46] F. Welsford-Ackroyd, A. Chalmers, R. K. dos Anjos, D. Medeiros, H. Kim, and T. Rhee. Asymmetric interaction between hmd wearers and spectators with a large display. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 670–671. IEEE, 2020. doi: 10.1109/VRW50115.2020.00186 2

[47] N. Wenk, J. Penalver-Andres, K. Buetler, T. Nef, R. M. Müri, and L. Marchal-Crespo. Effect of immersive visualization technologies on cognitive load, motivation, usability, and embodiment. *Virtual Reality*, pp. 1–25, 2021. doi: 10.1007/s10055-021-00565-8 2

[48] C. Wienrich and J. Gramlich. Appraisevr–an evaluation framework for immersive experiences. *i-com*, 19(2):103–121, 2020. doi: 10.1515/icom-2020-0008 2

[49] A. Winkler, J. Won, and Y. Ye. Questsim: Human motion tracking from sparse sensors with simulated avatars. In *SIGGRAPH Asia 2022 Conference Papers*, pp. 1–8, 2022. doi: 10.1145/3550469.3555411 3

[50] B. G. Witmer and M. J. Singer. Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, 06 1998. doi: 10.1162/105474698565686 3

[51] E. Wolf, D. Mal, V. Frohnapfel, N. Döllinger, S. Wenninger, M. Botsch, M. E. Latoschik, and C. Wienrich. Plausibility and perception of personalized virtual humans between virtual and augmented reality. In *2022 IEEE international symposium on mixed and augmented reality (ISMAR)*, pp. 489–498. IEEE, 2022. doi: 10.1109/ISMAR55827.2022.00065 2

[52] A. Yassien, P. ElAgroudy, E. Makled, and S. Abdennadher. A design space for social presence in vr. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society*, pp. 1–12, 2020. doi: 10.1145/3419249.3420112 1, 2

[53] B. Yoon, H.-i. Kim, G. A. Lee, M. Billinghurst, and W. Woo. The effect of avatar appearance on social presence in an augmented reality remote collaboration. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 547–556. IEEE, 2019. doi: 10.1109/VR.2019.8797719 2