# Exploring User Behaviour in Asymmetric Collaborative Mixed Reality

Nels Numan
nels.numan@ucl.ac.uk
University College London
London, United Kingdom

Anthony Steed
a.steed@ucl.ac.uk
University College London
London, United Kingdom

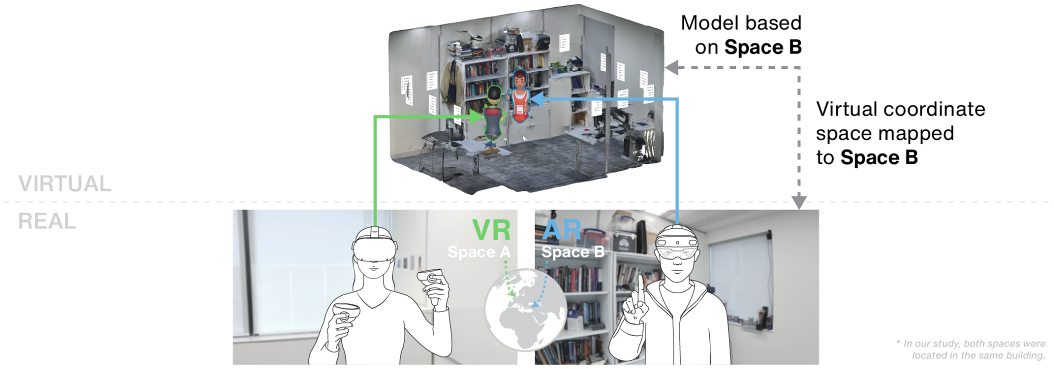

**Figure 1: Conceptual overview of our collaborative MR system, in which an AR and a VR user can collaborate remotely. The CVE contains a static 3D model of *Space B*. Users are represented with avatars. Line illustrations by Suhyun Park (artist).**

## ABSTRACT

A common issue for collaborative mixed reality is the asymmetry of interaction with the shared virtual environment. For example, an augmented reality (AR) user might use one type of head-mounted display (HMD) in a physical environment, while a virtual reality (VR) user might wear a different type of HMD and see a virtual model of that physical environment. To explore the effects of such asymmetric interfaces on collaboration we present a study that investigates the behaviour of dyads performing a word puzzle task where one uses AR and the other VR. We examined the collaborative process through questionnaires and behavioural measures based on positional and audio data. We identified relationships between presence and co-presence, accord and co-presence, leadership and talkativeness, head rotation velocity and leadership, and head rotation velocity and talkativeness. We did not find that AR or VR biased subjective responses, though there were interesting behavioural differences: AR users spoke more words, AR users had a higher median head rotation velocity, and VR users travelled further.

## CCS CONCEPTS

- **Human-centered computing → Mixed / augmented reality**; *Virtual reality*; **Collaborative interaction**.

## KEYWORDS

mixed reality, virtual reality, augmented reality, collaboration

**ACM Reference Format:**
Nels Numan and Anthony Steed. 2022. Exploring User Behaviour in Asymmetric Collaborative Mixed Reality. In *28th ACM Symposium on Virtual Reality Software and Technology (VRST '22), November 29-December 1, 2022, Tsukuba, Japan.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3562939.3565630

## 1 INTRODUCTION

An increasing amount of synchronous collaborative work takes place through video teleconferencing systems. However, most of these systems have severe difficulties in reproducing communicative interactions such as eye gaze and spatial references, which occur naturally in face-to-face settings [13]. The paradigm of mixed-reality (MR) technology, including virtual reality (VR) and augmented reality (AR), provides solutions to build systems that resolve these limitations due to its natural interface and the potential to be integrated with the real world. The potential of MR has prompted extensive research on using MR to replicate and augment various elements of face-to-face collaboration [7, 40].

Over the years, the field of collaborative MR has come a long way, with technological advancements and industry interest enabling the creation of impressive complex systems such as HoloPortation [30] and commercial products such as Horizon Workrooms [25]. However, little is known about how different types of MR devices can be used collaboratively. Collaborative MR with asymmetric interfaces – where users use different types of devices to collaborate (e.g., AR and VR) – poses novel challenges to creating CVEs where remote and local users can effortlessly work together. Examples of design challenges include enabling users to make spatial references and to

---

*VRST '22, November 29-December 1, 2022, Tsukuba, Japan*

ACM ISBN 978-1-4503-9889-3/22/11…$15.00
https://doi.org/10.1145/3562939.3565630

have a sense of awareness of other user's interaction with the CVE (i.e., workspace awareness [10]), which are commonly associated with increased task performance [10, 33, 34]. Moreover, it is unclear how asymmetry affects user experience and collaborative processes [3, 52]. Alleviating these barriers is of paramount importance not only for productivity and ease of use, but also to ensure every user can collaborate effectively regardless of the devices they are able to access. Furthermore, understanding the complexity of interaction between asymmetric interfaces could improve or enable collaborative scenarios that are inherently asymmetric, such as remote assistance [36, 45, 54] or training scenarios [39, 51]. For a broader overview of relevant applications, we refer to [7, 40].

Despite growing efforts, the mechanisms that underpin effective collaboration in asymmetric scenarios are not fully understood. Specifically, it is unclear to what extent interface asymmetry influences the collaborative process in collaborative MR. To address this, we present an exploratory study centred around the question:

Research Question 1. *Does asymmetry of interfaces bias interaction in the collaboration between an AR and a VR user?*

We investigated leadership emergence as one potential bias introduced by asymmetric interfaces. By studying leadership and other user behaviour, previous works identified proof indicating that leadership is conferred by immersion [1, 31, 43, 47], suggesting the existence of a bias introduced by interface asymmetry.

While arguments exist for either AR or VR being the more immersive interface, we base our first hypothesis on a study on AR-to-VR collaboration by Pan et al. [31], who found that AR users tend to emerge as the leader in dyadic collaboration. Furthermore, the aforementioned works considered talkativeness as an indicator of leadership, on which we base our second hypothesis.

Hypothesis 1. *AR users obtain a higher participant-rated leadership score than VR users.*

Hypothesis 2. *AR users obtain a higher participant-rated talkativeness score and speak more words than VR users.*

Furthermore, as VR users of our system have an easier way of navigating around the CVE using a joystick in addition to physical movement, we additionally hypothesise the following:

Hypothesis 3. *VR users travel further than AR users.*

To test our hypotheses, we conducted an exploratory experiment where dyads (i.e., pairs) of users used a VR HMD and an AR HMD to collaboratively perform a task. We employed an existing word puzzle task [43, 47], as it requires joint orientation, exploration, and problem-solving, which are aspects that are relevant to real-world applications such as inspection and training. We performed our analysis based on questionnaire responses and a set of behavioural measures that draw upon data that was recorded throughout the experimental trials. The results reveal no support for Hypothesis 1, but did partially support Hypothesis 2 and fully support Hypothesis 3. Furthermore, through post hoc analysis of the data, we identified positive relationships between presence and co-presence, accord and co-presence, leadership and talkativeness, head rotation velocity and leadership, and head rotation velocity and talkativeness. As such, our study provides further insight into how asymmetric interfaces in collaborative MR systems affect user behaviour.

## 2 RELATED WORK

In recent years, interest in asymmetric collaborative MR has been growing, as highlighted in several reviews on collaborative MR [7, 40]. In this section, we give a brief overview of related research on remote asymmetric collaborative MR.

Several early works have looked into the issue of effective and enjoyable collaboration between MR and less immersive devices [1, 3, 16, 43, 47]. In the first of a series of studies on the evaluation of the behaviour of small groups in this type of asymmetric scenario, Slater et al. [43] ran a study in which two desktop users and one VR HMD user performed a collaborative task. In this task, participants were assembled in a CVE and were asked to solve a collection of word puzzles. They found that the HMD user tended to emerge as the leader, indicating that immersion confers leadership. Furthermore, they found that presence and co-presence had a positive relation and that group accord increased with presence and task performance. In a subsequent study, Steed et al. [47] found further evidence to support these claims. Our experiment is based on the task used in this series of experiments [43, 47].

Axelsson et al. [1] performed a similar type of study in which a single desktop user collaborated with a single user of a CAVE-like VR system. The employed task required the manipulation of objects. The results of the study showed that immersed participants contributed to the task significantly more compared to less immersed users. However, there was no significant difference in verbal contribution among dyads.

More recently, attention has shifted towards collaboration between several types of MR interfaces, such as VR and AR. Pan et al. [31] studied collaboration with different combinations of MR devices: AR-to-AR, AR-to-VR, and AR-to-Desktop, where users collaborated on the design of a planet. In line with the aforementioned study, the authors found that immersion confers leadership. Interestingly, they identified that this effect occurs for 3D interactions (e.g., creation, deletion) but not for 2D interactions (e.g., modification of 2D surface properties).

Piumsomboon et al. [32] evaluated the effects of different view awareness cues for collaboration between a VR user and an AR user, such as eye gaze, head gaze, and field-of-view (FoV) frustum. The evaluated system was based on *CoVAR* [35] which, like our system, used a 3D reconstruction of the AR user's physical environment. In a user study, it was found that the proposed awareness cues improved task performance and usability. However, the system did not implement spatialized audio, which could have influenced the results. Furthermore, the authors observed that VR users performed better than AR users and took a more proactive role throughout the task, indicating a potential leadership effect. However, no formal evaluation of leadership was carried out.

Another large body of work investigated the representation of physical environments of local users through 360° video [20, 28, 33, 34, 49], which is an effective solution when a real-time view of an environment is needed but permits limited depth perception and control over the viewpoint of remote users. Teo et al. [50] combined 360° video and 3D models and found that users prefer to be able to switch between both for a collaborative search task.

Other works looked into asymmetry of interfaces in combination with robotic MR telepresence [14, 48]. For example, *VROOM* [14], a system that enabled remote VR users to be locally represented by a telepresence robot equipped with a 360° camera, which was overlaid with an avatar of the remote VR user for local AR users. In the design of the system, five common issues of asymmetry were considered: embodiment, expressiveness, mobility, awareness, and presence. A user study revealed important insights into the challenges of asymmetric telepresence systems, including avatar representations and uncertainties regarding others' capabilities and viewing direction. Some works specifically looked into some of these issues, such as methods for gaze and gesture cues [19–22, 32, 33, 55], and avatar representation [56].

Overall, these works highlight the breadth of the issues caused by asymmetry, with many questions remaining unanswered. This exploratory study aims to further our understanding of how users collaborate with asymmetric MR interfaces.

## 3 SYSTEM DESIGN

Our system was built on Ubiq [8], which is an open-source social VR framework for research. As Ubiq is currently only available for Unity, our system was implemented in Unity (version 2020.6.26f1).

Our system enables asymmetric collaboration between an AR user and a VR user. Both users were embodied as a random avatar from the Ubiq avatar module. These avatars are cartoon-like, have simple shapes, and only have a head, upper torso, and floating hands. The floating hands have a mitten-like shape and include an animation that closes the hand. Avatar appearance was device-specific and remained unchanged throughout each trial. Avatar movement was networked using the networking module of Ubiq, which went through a local network that both devices were connected to through Wi-Fi. Both users were connected through a spatialized audio connection (stereo with attenuation) established through voice over IP (VoIP) using Ubiq.

*Apparatus.* For the AR user, we used a Microsoft HoloLens 2 (HL2). Hand tracking was implemented to allow AR users to control their avatar's hands with their own hands. Mixed Reality Toolkit [26] (MRTK) was used for hand tracking and configuration of the system. For the VR user, we used a Meta Quest 2 (MQ2). Avatar hands were controlled with the MQ2 controllers. Interactions were controlled by Ubiq's interactions module. Both users had control over the hand-closing animation.

*Room Model.* Dyads collaborated in a CVE containing a static 3D model of the physical space of the AR user. This 3D model was created with a fifth-generation 12.9-inch iPad Pro running iOS 15 equipped with a LiDAR, using an application called *3D Scanner App* [17]. Post-processing of the model was done in Blender [6].

*Coordinate Space.* Both users had an egocentric point of view. They were scaled to their normal height and could move independently about the CVE. They were thus free to choose how to stand relative to their companion.

To establish a shared coordinate system between the AR user and VR user, the CVE was aligned with the physical room of the AR user. We used a manual alignment procedure based on two known points $Q1$ and $Q2$, located on two corners of a large table, as described by McGill et al. [24]. Upon the first launch of the system on the HL2, an initial calibration was needed. This was done by dragging two spheres to their known location on the corners of the large table in the room. *Space pins* from the world-locking tools module from MRTK were set once the initial setup was completed to lock the virtual coordinate space to the real world. The calibration persisted across sessions.

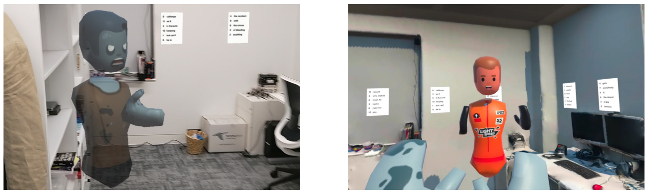

**Figure 2: Screenshots of an AR (left) and VR (right) user's perspective. *Note: these images are not temporally aligned.***

*Experimental Space.* Dyads carried out the task in two separate spaces on the same floor of a university building. The VR user and the AR user were located in *Space A* and *Space B*, respectively, as shown in Figure 1. For the AR user, virtual posters were placed along the walls. These posters were placed in corresponding locations in the 3D model for the VR user. The 3D model of the room was only visible to the VR user. Perspectives of an AR and VR user are shown in Figure 2.

*Recordings.* The record and replay module from *Ubiq-Exp* [46] was integrated with the system. This module recorded all network messages in addition to microphone audio, which allowed us to perform post hoc analysis of the trials including the extraction of additional metrics even after trials were completed.

*Interactions.* Users could freely move through the space but could not interact with any virtual objects. Both users could use their avatar's hands for simple gestures. The AR user could move through the physical room by walking around physically, whereas the VR user could move around with the joystick on their left controller or could walk around within the boundaries of their physical space. The VR user could not collide with any objects in the virtual room.

## 4 METHOD

### 4.1 Design

Our study followed a within-subjects design, where the order of the conditions was counterbalanced using a Latin square. We assessed the effect of one independent variable, which is the HMD type used during the task: a VR HMD (MQ2) or an AR HMD (HL2). As such, each participant within a dyad performed the task using each HMD by switching after the first trial period.

### 4.2 Participants

We recruited fifteen dyads, resulting in 30 participants in total. Participants were recruited through mailing lists and posters spread across a university campus. Participants were in the range of 18-24 years old ($N = 21$) or 25-34 years old ($N = 9$), with 15 women and 15 men. Eleven of the fifteen dyads knew each other prior to the experiment. Most participants did not speak English natively ($N = 18$). Furthermore, half of the participants played (online) video games weekly ($N = 15$), and some participants used VR ($N = 7$) or social VR ($N = 2$) weekly.

## 4.3 Task

The task design of this experiment was analogous to the work of Slater et al. [43] and Steed et al. [47], but with dyads instead of groups of three. The word puzzle task involved identifying and ordering same-numbered words across virtual posters hung around a (virtual) room. Each poster contained a list of five to six words, preceded by a number that indicated which of the eleven riddles it belonged to. Participants were asked to solve the riddles by finding all same-numbered words and putting them in the right order. Participants could check their solution at any time by saying it out loud, after which the experimenter told them whether it was correct. The model and posters used are available at https://osf.io/fwru9/.

## 4.4 Procedure

The study was led by two experimenters: the main experimenter and an assistant experimenter. Before the experiment, participants were asked to read through an information sheet and fill out a pre-study questionnaire. Each participant was assigned to one of two conditions at random for the first trial period. Participants switched conditions for the second trial period. Each period took 15 minutes.

Before the first trial period, each participant was led to their assigned space and was instructed on how to use their HMD by an experimenter. Once both participants were instructed, the main experimenter took place in front of a PC in a separate space, running a desktop client of the system. Represented as a random avatar, the main experimenter then proceeded to verbally explain the task to the participants. Once the task was clear to the participants, the experimenter's avatar disappeared and the trial started.

Once the first trial period ended, both participants were assisted in taking off their HMDs and were led to desks where they filled out a post-study questionnaire with a PC. Once finished, they received new instructions and performed the task under the other condition. After the second trial, participants filled out the same questionnaire.

After each trial, an unstructured interview was conducted to learn more about the participants' experiences and preferences. The participants received a £10 voucher as compensation.

## 4.5 Measures

### 4.5.1 *Pre-trial Questionnaire.*

The pre-trial questionnaire consisted of two sections. Firstly, a section on demographics including age, gender, and experience with video games and MR. The second section contained the Interaction Anxiousness Scale (IAS) questionnaire [18]. Following studies that we draw upon [43, 47], this questionnaire was added to obtain a measure of social anxiety, which is theorised to influence participants' leadership capabilities. We also asked participants to fill in the BFI-10 questionnaire [37], which measures personality in terms of five dimensions that have been shown to be associated with leadership capability [15, 41]. However, the results from the latter questionnaire were not considered in the present study, as our sample size was too small to jointly analyse the five dimensions of the scale without the risk of overfitting statistical models. A 5-point Likert scale was used for both questionnaires.

### 4.5.2 *Post-trial Questionnaire.*

After each trial period, participants filled in a post-trial questionnaire. The statements of the post-trial questionnaire are shown in this article's supplementary material (Table S1). These statements are based on previous work that measured leadership emergence in asymmetric collaborative MR [31, 43, 47]. We employed the following measures: talkativeness, leadership, presence, co-presence, and accord (i.e., *group harmony*). As most dyads had a prior relationship, we excluded two irrelevant accord-related questions.

Given the absence of an established questionnaire for measuring presence across the spectrum of MR, we used a modified version of the presence-related questions from [43, 47]. Specifically, we prepended "virtual" to the word "office" to differentiate between the virtual and real space.

To measure leadership, each participant was asked to rate their own and their partner's degree of leadership in carrying out the task on a scale from 1 to 100. Furthermore, on the same scale, each participant rated their own and their partner's talkativeness. Ratios were calculated based on the given scores to obtain leadership scores and talkativeness scores that both added up to 1. The final participant-rated leadership and talkativeness scores that were used in the analysis were based on the average of the participant's self-rated score and the score received from the other participant.

The degree of presence and co-presence were rated by participants on a 7-point Likert scale, where '7' represented a high degree of the respective measure. The number of '6' and '7' responses were counted to form presence and co-presence scores. Accord was also rated on a 7-point Likert scale, where '7' represents a high degree of accord. Accord scores were defined as the average of responses.

### 4.5.3 *Behavioural Measures.*

Motion and audio data of the participants were collected during the trial. For post hoc analysis, behavioural measures were defined based on the recorded data: number of words spoken, distance travelled, and median head rotation velocity. These measures provide insight into user's level of activity, which could reveal the impact of the MR interface on user behaviour. (1) *Number of words spoken* was calculated by transcribing the speech of each participant and counting the resulting words. This measure gave us an objective method of measuring talkativeness, which previously has been associated with leadership [43, 47]. (2) *Distance travelled* was calculated based on the distance that each participant's avatar travelled within the CVE. This measure gave us an objective method of measuring user activity. (3) *Median head rotation velocity* was calculated based on avatar head movement driven by HMD sensors. We attempted to mitigate sensor noise by applying a moving average filter to the data (recorded at 30Hz). Furthermore, the signal of the VR HMD contained dropped frames, which we removed by linearly interpolating between neighbouring values. An example of the filtered signals is shown in this article's supplementary material (Figure S1). For analysis, we took the median of the filtered signals of the pitch, roll, and yaw axes.

The optimal sliding window size to recover the true signal is unknown, and visual inspection of the signal provided limited insight. Therefore, we performed our statistical analysis on different window sizes, to attempt to account for a potential bias introduced by the filter. Our main findings are based on the window size of 3 frames, which we selected based on visual inspection of plots and recordings. We report statistics on larger window sizes (5, 10, and 15 frames) in this article's supplementary material (Table S2–S5).

## 5 RESULTS

In this section, we first report results that highlight the differences across HMD types, including the results relevant to our hypotheses. Next, we present results that highlight differences across the two trial periods. Lastly, we report notable relations between measures.

As we used modified versions of existing questionnaires, we assessed the reliability of the questionnaire with Cronbach's alpha. The results showed that the co-presence subscale had good reliability (0.85 for eight items), whereas the presence (0.68 for six items) and accord (0.69 for five items) subscales had acceptable reliability for usage in exploratory research [11].

Some data was excluded from our analysis. Firstly, behavioural measures based on recorded data are missing for trial 10 due to a system error. Secondly, the responses to accord-related questions of one of the participants from trial 6 have been excluded as they all had the value of '1' and were the last questions of the questionnaire.

### 5.1 Differences Across HMD Types

Paired-samples two-tailed t-tests were used to assess whether there were statistically significant differences between the mean of questionnaire responses and behavioural measures of users when using an AR HMD or VR HMD.

Based on boxplots, outliers were detected that were more than 1.5 box lengths from the edge of the box. However, with exception of the excluded outliers mentioned above, inspection of their values did not reveal them to be extreme and they were kept in the analysis. The assumption of normality was not violated for any of the variables, as assessed by Shapiro-Wilk tests. The means, standard deviations, and statistical test results for each variable are shown in Table 1. Our key observations are as follows.

- We found *no* significant differences in the values of subjective measure between AR and VR users, including leadership and talkativeness scores.
- AR users spoke a significantly higher number of words compared to VR users ($p = .020$).
- VR users travelled a significantly further distance than AR users ($p = .036$).
- AR users had a significantly higher median head rotation velocity than VR users in the yaw ($p = .009$), roll ($p < .001$), and pitch axis ($p < .001$), indicating that AR users turned their heads more often or with a higher velocity.

Results of differences between HMD types per trial period are shown in this article's supplementary material (Table S6). In this table, we can observe that the difference in means of median head rotation was only significant in the first trial period. Furthermore, while the difference in the mean distance travelled was present in both trial periods, it was only significant for the second trial period.

Notably, no significant difference was found between the leadership scores of AR and VR users. Following Steed et al. [47], we additionally performed a regression analysis to assess the relationship between the participant IAS and leadership across HMD types. Since our leadership scores are a ratio within the range of 0 and 1, we performed a beta regression analysis (with R package `betareg` [57]). A model was constructed for each trial period, immersion level, and participant sex. The resulting model equations are shown in Table 2, based on which it can be noted that the coefficients of the IAS were only significant for females who used the AR HMD first or the VR HMD second.

### 5.2 Differences Across Trial Periods

Paired-samples two-tailed t-tests were used to assess whether there were statistically significant differences between the mean of questionnaire responses and behavioural measures between the first and second trial periods.

Based on boxplots, outliers were detected that were more than 1.5 box lengths from the edge of the box. However, with exception of the excluded outliers mentioned above, inspection of their values did not reveal them to be extreme and they were kept in the analysis. The assumption of normality was not violated for any of the variables, as assessed by Shapiro-Wilk tests.

The means, standard deviations, and statistical test results for each variable are shown in Table 3. In this table, it can be observed that solely the mean of the median head rotation in the yaw axis was significantly higher in the second trial period compared to the first trial period ($p = .048$). Results of the difference between trial periods per HMD type are shown in this article's supplementary material (Table S7). Based on this table, it can be noted that the difference in mean of the median head rotation (in all three axes) for the second trial period was present solely for the VR condition.

### 5.3 Relations Between Measures

In this section, we report on notable relations between variables that were identified in our statistical analysis. Where relevant, we describe relations per HMD type and per trial period.

*5.3.1 Presence and Co-presence.* A Spearman's rank-order correlation test was run to assess the relationship between presence and co-presence for each HMD type and trial period. Preliminary analysis showed these relationships to be monotonic, as assessed by visual inspection of scatter plots. There was a statistically significant strong positive correlation for both the AR condition ($r_s(28) = .471$, $p = .009$) and VR condition ($r_s(28) = .685$, $p < .001$). Upon evaluating this relation among the trial periods, we found a statistically significant strong positive correlation for the first period ($r_s(28) = .647$, $p < .001$) and a statistically significant moderate positive correlation for the second period ($r_s(28) = .555$, $p = .001$).

*5.3.2 Accord and Co-presence.* A Spearman's rank-order correlation test was run to assess the relationship between accord and co-presence for each HMD type and trial period. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. There was a statistically significant strong positive correlation for the AR condition ($r_s(27) = .603$, $p < .001$) and VR condition ($r_s(27) = .419$, $p = .020$). Furthermore, there was a statistically significant strong positive correlation for the first trial period ($r_s(27) = .760$, $p < .001$) but not for the second trial period ($r_s(28) = .330$, $p = .075$).

*5.3.3 Leadership and Talkativeness.* A Spearman's rank-order correlation test was run to assess the relationship between participant-rated leadership and participant-rated talkativeness for each HMD

**Table 1: Results of paired t-test for all variables between HMD conditions. Significance levels: $^{*}p < .05$, $^{**}p < .01$ $^{***}p < .001$.**

| HMD | AR | | VR | | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | M | SD | M | SD | df | t | p | Cohen's d |
| Presence | 2.47 | 1.85 | 2.37 | 1.81 | 29 | 0.2 | 0.843 | 0.06 |
| Co-presence | 3.13 | 2.74 | 3.73 | 2.24 | 29 | -1 | 0.327 | -0.24 |
| Accord | 5.84 | 0.89 | 5.91 | 0.91 | 29 | -0.5 | 0.620 | -0.08 |
| Leadership | 0.51 | 0.08 | 0.49 | 0.08 | 29 | 0.89 | 0.381 | 0.15 |
| Talkativeness | 0.5 | 0.08 | 0.5 | 0.08 | 29 | 0.06 | 0.954 | 0.01 |
| Words spoken | 854.63 | 280.36 | 782.93 | 286.03 | 25 | 2.49 | $0.020^{*}$ | 0.25 |
| Distance travelled | 144.09 | 53.69 | 178.92 | 88.69 | 27 | -2.2 | $0.036^{*}$ | -0.47 |
| Median yaw | 0.23 | 0.05 | 0.2 | 0.06 | 27 | 2.83 | $0.009^{**}$ | 0.55 |
| Median roll | 0.07 | 0.02 | 0.05 | 0.02 | 27 | 4.77 | $< 0.001^{***}$ | 0.77 |
| Median pitch | 0.08 | 0.02 | 0.06 | 0.02 | 27 | 5.61 | $< 0.001^{***}$ | 0.81 |

**Table 2: Regression of leadership score for both trial periods. Significance indicated with $^{*}p < .05$, $^{**}p < .01$ $^{***}p < .001$.**

| | | Participant sex | |
|---|---|---|---|
| Period | HMD | Female | Male |
| 1 | AR | $2.125 - 0.052^{*} \times IAS$ | $0.191 - 0.004 \times IAS$ |
| | VR | $-0.390 + 0.014 \times IAS$ | $0.130 - 0.008 \times IAS$ |
| 2 | AR | $0.034 + 0.001 \times IAS$ | $0.121 - 0.005 \times IAS$ |
| | VR | $1.934 - 0.051^{***} \times IAS$ | $-0.031 + 0.003 \times IAS$ |

type and trial period. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. There was a statistically significant strong positive correlation for the AR condition ($r_s(28) = .764, p < .001$) and VR condition ($r_s(28) = .764, p < .001$). We note that these correlations are equal, as the leadership and talkativeness scores are ratios that are symmetric per dyad. Furthermore, there was a statistically significant strong positive correlation for the first trial period ($r_s(28) = .952$, $p < .001$) and second trial period ($r_s(28) = .605, p < .001$).

In addition, a Spearman's rank-order correlation test was run to assess the relationship between participant-rated leadership and the number of words spoken by each participant for each HMD type and trial period. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. There was no statistically significant correlation for the VR condition ($r_s(28) = .199, p = .320$) or AR condition ($r_s(28) = .374, p = .054$). However, there was a statistically significant strong positive correlation for the first trial period ($r_s(28) = .490, p = .011$), but not for the second trial period ($r_s(28) = .200, p = .308$).

*5.3.4 Head Rotation Velocity and Leadership.* A Spearman's correlation test was run to assess the relationship between participant-rated leadership and head rotation velocity for each HMD type and trial period. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. For the AR condition, there was a statistically significant strong positive correlation between participant-rated leadership and median yaw ($r_s(26) = .492, p = .008$) and median roll ($r_s(26) = .466$, $p = .012$), but not for median pitch ($r_s(26) = .305, p = .114$). For the VR condition, there was a statistically significant strong positive correlation between participant-rated leadership and median yaw ($r_s(26) = .458, p = .014$), but not for median roll ($r_s(26) = .283$, $p = .144$), and median pitch ($r_s(26) = .291, p = .134$). For the first trial period, there was a statistically strong significant positive correlation between participant-rated leadership and median yaw ($r_s(26) = .500, p = .007$), median roll ($r_s(26) = .496, p = .007$), and median pitch ($r_s(26) = .410, p = .030$). For the second trial period, there was a statistically significant strong positive correlation between participant-rated leadership and median yaw ($r_s(26) = .495$, $p = .007$), but not for median roll ($r_s(26) = .300, p = .121$) and median pitch ($r_s(26) = .128, p = .517$).

*5.3.5 Head Rotation Velocity and Talkativeness.* A Spearman's correlation test was run to assess the relationship between participant-rated talkativeness and head rotation velocity for each HMD type and trial period. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. For the AR condition, there was a statistically significant moderate positive correlation between participant-rated talkativeness and median yaw ($r_s(26) = .397, p = .036$), but not for median roll ($r_s(26) = .340, p = .077$) and median pitch ($r_s(26) = .225, p = .250$). For the VR condition, there was a statistically significant strong positive correlation between participant-rated talkativeness and median yaw ($r_s(26) = .445, p = .018$), but not for median roll ($r_s(26) = .327, p = .090$) and median pitch ($r_s(26) = .220, p = .262$). For the first trial period, there was a statistically significant strong positive correlation between participant-rated talkativeness and median yaw ($r_s(26) = .503, p = .006$), median roll ($r_s(26) = .486$, $p = .009$), and median pitch ($r_s(26) = .378, p = .048$). For the second trial period, there was no statistically significant positive correlation between participant-rated talkativeness and median yaw ($r_s(26) = .284, p = .143$), median roll ($r_s(26) = .082, p = .680$), and median pitch ($r_s(26) = -.094, p = .634$).

In addition, a Spearman's correlation test was run to assess the relationship between the number of words spoken and head rotation velocity for each HMD type and trial period. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. For the AR condition, there was

**Table 3: Results of paired t-test for all variables between trial periods. Significance levels: $^{*}p < .05$, $^{**}p < .01$ $^{***}p < .001$.**

| Period | 1 | | 2 | | | | | |
|---|---|---|---|---|---|---|---|---|
| Variable | M | SD | M | SD | df | t | p | Cohen's d |
| Presence | 2.03 | 1.47 | 2.8 | 2.06 | 29 | -1.59 | 0.122 | -0.42 |
| Co-presence | 3.87 | 2.52 | 3 | 2.45 | 29 | 1.47 | 0.153 | 0.35 |
| Accord | 5.82 | 0.99 | 5.93 | 0.81 | 29 | -0.78 | 0.442 | -0.13 |
| Leadership | 0.5 | 0.09 | 0.5 | 0.07 | 29 | 0 | 1 | 0 |
| Talkativeness | 0.5 | 0.11 | 0.5 | 0.06 | 29 | 0 | 1 | 0 |
| Words spoken | 832.65 | 267.27 | 805.89 | 300.9 | 25 | 0.55 | 0.585 | 0.09 |
| Distance travelled | 168.61 | 87.3 | 154.41 | 60.44 | 27 | 0.84 | 0.410 | 0.19 |
| Median yaw | 0.2 | 0.06 | 0.22 | 0.05 | 27 | -2.07 | 0.048* | -0.42 |
| Median roll | 0.06 | 0.03 | 0.07 | 0.02 | 27 | -1.27 | 0.213 | -0.27 |
| Median pitch | 0.07 | 0.03 | 0.08 | 0.02 | 27 | -1.65 | 0.110 | -0.34 |

a statistically significant positive correlation between participant-rated leadership and median yaw ($r_s(26) = .487$, $p = .011$), median roll ($r_s(26) = .398$, $p = .041$), and median pitch ($r_s(26) = .421$, $p = .030$). For the VR condition, there was no statistically significant correlation between participant-rated leadership and median yaw ($r_s(26) = .295$, $p = .135$), median roll ($r_s(26) = .168$, $p = .401$), and median pitch ($r_s(26) = .292$, $p = .139$). For the first trial period, there was no statistically significant correlation between participant-rated leadership and median yaw ($r_s(26) = .342$, $p = .088$), median roll ($r_s(26) = .269$, $p = .183$), and median pitch ($r_s(26) = .320$, $p = .111$). For the second trial period, there was a statistically significant positive correlation between participant-rated leadership and median yaw ($r_s(26) = .515$, $p = .006$) and median pitch ($r_s(26) = .391$, $p = .041$), but not for median roll ($r_s(26) = .295$, $p = .128$).

*5.3.6 Number of Words Spoken and Participant-rated Talkativeness.* A Spearman's correlation test was run to assess the relationship between participant-rated talkativeness and the number of words spoken. Preliminary analysis showed these relationships to be monotonic as assessed by visual inspection of scatter plots. There was a statistically significant moderate positive correlation for the AR condition ($r_s(26) = .385$, $p = .047$), but not for the VR condition ($r_s(26) = .179$, $p = .372$). There was a statistically significant strong positive correlation between participant-rated talkativeness and number of words spoken for the first trial period ($r_s(26) = .439$, $p = .025$), but not for the second trial period ($r_s(26) = .259$, $p = .183$).

*5.3.7 Accord and Task Performance.* A Spearman's correlation test was run to assess the relationship between dyad-level accord and task performance (number of solved riddles). Preliminary analysis showed the relationship to be monotonic for both trial periods, as assessed by visual inspection of scatter plots. There was no statistically significant correlation between dyad-level accord and task performance for the first trial period ($r_s(26) = -.364$, $p = .057$) or the second trial period ($r_s(28) = 0.236$, $p = .209$).

## 6 DISCUSSION

### 6.1 Insights from the Statistical Analysis

*6.1.1 Leadership and Talkativeness.* We did not find sufficient evidence to support Hypothesis 1, as the results in Section 5.1 did not show a significant difference between participant-rated leadership scores of AR and VR users. This may imply that the difference in immersion level between these types of devices is not large enough to make a significant impact on leadership emergence. Specifically, both users wore an HMD and had roughly the same capabilities.

However, it is conceivable that this is a consequence of different factors. Firstly, whereas previous works [43, 47] studying leadership emergence in asymmetric collaboration studied the behaviour of groups of three users, we studied collaboration between two users. Dyadic collaboration possibly needs less coordination due to the smaller size of the group and direct interdependence among collaborators. Therefore, future work may assess the impact of group size on leadership emergence in collaborative MR. Secondly, the social distance between dyads may have played a role [29]. Lastly, while the IAS of participants could have played a role, the presented regression equations in Table 2 provide little insight due to their mixed and largely statistically insignificant results.

On the other hand, we found partial support for Hypothesis 2. Similar to participant-rated leadership, the results do not show a significant difference between participant-rated talkativeness scores of AR and VR users. However, the mean number of words spoken did turn out significantly higher for AR users when compared to VR users. With this, we identified objective evidence to support Hypothesis 2, but lack support from the subjective measure. The disagreement between these measures is highlighted by the fact that they were only significantly positively related for the first trial period, as shown in Section 5.3.3. This discrepancy raises questions regarding the optimal choice of measures and highlights the importance of combining subjective measures with objective ones.

*6.1.2 Presence, Co-presence, and Accord.* As mentioned in Section 4.5.2, it is unclear whether the employed questionnaire provides meaningful insight into the sense of presence and co-presence of AR users. Therefore, the presented results pertaining to these factors should be interpreted with caution. Nevertheless, based on our statistical results presented in Table 1, we found no significant difference in presence scores between the AR and VR HMD. Similarly, we did not find a statistically significant difference between co-presence scores between AR and VR.

Presence and co-presence are conceptually independent. However, there is contradictory evidence as to whether co-presence and presence are related [4]. The results of our statistical analysis in Section 5.3.1 suggest a strong positive correlation between the two, which aligns with previous studies with similar experimental designs to ours [1, 43, 47]. However, causal factors leading to this relationship remain unclear. Nonetheless, keeping in mind the aforementioned uncertainty of the reliability of the questionnaire, increasing presence appears to potentially result in an increase in co-presence or vice-versa.

We also found co-presence positively correlated with accord for the first session, as reported in Section 5.3.2, which corresponds with previous findings [43, 47]. While we do not have enough information to establish the cause of this relationship, we theorise that a stronger sense of *being together* may intensify the sense of harmony with the other, or vice versa. However, we did not identify this relationship for the second trial period, signifying that the relationship may fade once users have acclimatised with the workings of the system, their collaborator, or the task.

Throughout each trial period, participants collaborated with the same person. Therefore, we expected accord to grow over time. Table 1 reveals a negligible increase of accord in the second period over the first period, which was not statistically significant. This may be influenced by the fact that the majority of dyads knew each other prior to the experiment.

As opposed to what was found in previous work [43, 47], we did not find a positive relation between participant- or dyad-level accord and task performance based on the results presented in Section 5.3.7. This also may be related to prior social relationships of the dyads, as enjoyment of collaboration is less likely to depend on task performance, which possibly was the case to a higher degree in previous studies where participants did not know each other yet.

Further research is needed to explore the importance of presence, co-presence, and accord on task performance.

*6.1.3 Head Rotation and Distance Travelled.* The results presented in Table 1 show that VR users travelled a significantly longer distance than AR users. This supports our hypothesis that VR users move around the CVE more than AR users. We theorise that this is a consequence of their unique capability to manoeuvre using the VR controller's joystick. Further, several participants mentioned better readability of posters through the AR HMD compared to the VR HMD, which may have prompted VR users to move around more to position themselves at ample reading distance. The AR user in the second period of trial 8 stated the following.

**AR** "I can see much better with this headset because last time I couldn't really read the posters unless I'm really close to it."

Moreover, in Table 1, it can be observed that AR users tend to have a higher median head rotation velocity compared to VR users. This aligns with the theory that, to compensate for moving around less than VR users, AR users may rotate their heads more to look at posters around the room. The most common rotation axis for this type of movement is the yaw axis, which correspondingly has the highest effect size, as noted in Table 1. Another explanation of this effect is the narrow FoV of the AR HMD, forcing users to rotate their head more often to observe objects and spatial references [12].

In the presented analysis in Section 5.3.4, we found that head rotation velocity in the yaw axis is related to leadership score. A possible cause for this is that leaders may be more likely to actively look around the CVE to gather information to solve the task. Similarly, head rotation velocity was found to be correlated with talkativeness for the first trial period, which may mean that users who look around more have more to talk about. However, this relation was not identified in the second trial period, so is unlikely to be generalisable.

## 6.2 Spatial Referencing

During the study, we observed that several dyads referred to elements of the environment to coordinate their conversation and collaboration. Similarly, some dyads started the task by establishing whether they were seeing the same virtual posters in the same location. In this process, dyads made use of elements around the office, read out words shown on the posters, or counted the posters. This highlights the importance of shared visual landmarks; which has been investigated by Müller et al. [27]. An example of such an interaction in the first period of trial 2 is shown below.

**AR** "I don't know if the posters are in the same positions, but I can see ‘except’."

**VR** "There's like a door, right? There's a door and a clock right above it."

**AR** "Yes, yeah, same. So it's the same."

On two occasions, dyads doubted whether the position or contents of the posters was changing over time but carried on with the task after a brief pause. In some cases, verbal spatial references resulted in confusion, as participants regularly were unaware of what the other participant was pointing at. In this case, it took dyads extra time to establish common ground. This indicates a shortcoming of our system [2], which only supports pointing with the avatar's hand. Another obstacle to spatial referencing was the limited FoV and brightness of the AR HMD. Future work may explore video see-through HMDs, which would allow for an evaluation of the AR and VR conditions using the same HMD. During the second period of trial 2, a participant commented:

**VR** "You're so much more visible to me this time around, whereas last time I had to focus to find you."

Existing work addressed these problems by adding different types of gestural or gaze cues, which have been shown to improve task performance and usability [19–22, 32, 33, 55]. Adding these types of cues to our system and study is a subject of future work.

## 6.3 Spatial Behaviour and Strategy

Motion data of the avatars of participants were recorded. Aside from several objective measures such as the total distance travelled and the median head rotation velocity, we visually evaluated the recorded head position to discover patterns or notable movement behaviour. In general, strategy did not appear to be influenced by trial period or HMD type. Further, VR users stayed within the bounds of the room and rarely moved through objects, although there were none that obstructed them from carrying out the task.

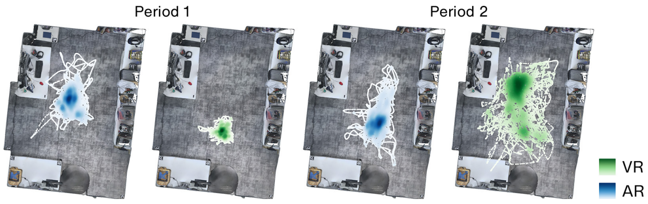

**Figure 3: Plot of head positions of each HMD type for the two periods of trial 14. Background is a top view of the CVE.**

While objects were generally avoided, participants often collided with each other. At times, this resulted in interruptions. Two participants were particularly aware of their collaborator bumping into them. This was exemplified by their tendency to tell their collaborator that they were obstructing them or apologise when they noticed they obstructed them, which aligns with findings of previous work [43]. These comments could not be linked to a specific HMD type.

In four trials, a *divide and conquer* strategy was verbally agreed upon by participants and was observable in plots of head position data. An example is shown in Figure 3, where it can be seen that participants primarily move about in specific parts of the room. In general, strategies nor movement patterns appeared to be related to trial period or HMD type.

## 6.4 Environmental Representation

For the VR user, the CVE contained a static 3D model of the AR user's physical space, as shown in Figure 1. Since participants used both HMD types in a randomised order, participants got to observe their collaborator's perspective in the second trial period, which resulted in notable reactions. For example, the majority of participants who first performed the task in VR did not immediately recognise they had entered the physical version of the CVE upon entering Space B while getting set up for the second trial period. In some cases, participants were unaware that their collaborator was located in the (virtual) space during the first trial period. At the start of the second period of trial 1, a participant shared their feeling of uncanniness after switching to VR.

**VR** "It's quite interesting doing it. When you are first in the real room and then in this one, which is a reproduction of that room. I feel like that changed my perspective."

**VR** "Cause now I'm sort of, you know, predisposed to the space already, but it's a slightly uncanny thing."

The issue of the effect of environmental representations in asymmetric collaborative MR is an intriguing one that could be explored in further research. In particular, the impact of model fidelity and dynamic updating mechanisms would be interesting to study.

## 6.5 Limitations

There are a number of limitations to the presented results. Firstly, our study did not include conditions for co-located or symmetric setups, which could have enabled insightful comparisons [9]. Secondly, as mentioned in Section 4.5.2, the reliability of the used presence and co-presence scales are uncertain, considering they were originally developed for VR, which is conceptually different from AR. While a few questionnaires for measuring presence in MR have been proposed [38, 53], measuring presence and co-presence in MR remains an open problem [42]. Thirdly, participants had verbal contact with the experimenter during the experiment, which could have led to breaks in presence [44]. Future work may implement other mechanisms to indicate solution correctness. Fourthly, switching between conditions and spaces could have influenced the results of the second trial period. For this reason, where relevant, we reported and discussed results for both periods separately. Lastly, participants were represented by random cartoon-like avatars, which could have impacted interaction [23]. Future research could explore the implications of avatar properties, including gender and fidelity.

## 7 CONCLUSION

The impact of asymmetry in collaborative MR has not been widely explored. The central aim of this work was to gain an understanding of how asymmetric interfaces in collaborative MR systems bias the behaviour of users. Towards this, we presented an exploratory study to investigate asymmetric collaboration between users of an AR and VR HMD. We performed our analysis based on questionnaire responses and behavioural measures that drew upon data that was recorded throughout the experimental trials. Following previous work [1, 31, 43, 47], we assessed whether AR users were more likely to arise as leaders compared to VR users. While we did not identify support from subjective responses for this, we did find that AR users objectively spoke more words than VR users, which has been considered as an indicator of leadership in existing works [43, 47].

In addition, with the aim of generating hypotheses for future work, we conducted a post hoc analysis. We found positive relationships between presence and co-presence, accord and co-presence, leadership and talkativeness, head rotation velocity and leadership, and head rotation velocity and talkativeness. Overall, the addition of objective measures proved revealing of user behaviour and supported the interpretation of subjective responses. In future research, we plan to explore measures based on eye gaze and motion data.

Another important issue for further research is to understand if differences in user behaviour among HMD types are essentially caused by restrictions or complications, or could be considered positive aspects enabling more effective collaboration (e.g., the ability to travel more quickly in VR). Moreover, we plan to investigate other asymmetric factors such as role asymmetry [3] and awareness [14]. In addition, areas to explore are collaboration over longer periods of time, with multiple task types, with various group sizes, and across different places to assess how these systems could support everyday collaborative work [5].

Altogether, our work provides additional insights over previous work as well as noteworthy observations that could serve as a base for future studies into asymmetric collaborative MR.

## 8 ETHICS STATEMENT

This study has been approved by the University College London Research Ethics Committee (Study ID number 4547/012).

## ACKNOWLEDGMENTS

This project has received funding from the European Union's Horizon 2020 Research and Innovation program under grant agreement No 739578.

# REFERENCES

[1] Ann-Sofie Axelsson, Åsa Abelin, Ilona Heldal, Ralph Schroeder, and Josef Wideström. 2001. Cubes in the Cube: A Comparison of a Puzzle-Solving Task in a Virtual and a Real Environment. *Cyberpsychology & behavior : the impact of the Internet, multimedia and virtual reality on behavior and society* 4 (May 2001), 279–86. https://doi.org/10.1089/109493101300117956

[2] Martin Bauer, Gerd Kortuem, and Zary Segall. 1999. " Where are you pointing at?" A study of remote collaboration in a wearable videoconference system. In *Digest of Papers. Third International Symposium on Wearable Computers*. IEEE, 151–158. https://doi.org/10.1109/ISWC.1999.806696

[3] M. Billinghurst, S. Bee, J. Bowskill, and H. Kato. 1999. Asymmetries in collaborative wearable interfaces. In *Digest of Papers. Third International Symposium on Wearable Computers*. 133–140. https://doi.org/10/dqx596

[4] Saniye Tugba Bulu. 2012. Place presence, social presence, co-presence, and satisfaction in virtual worlds. *Computers & Education* 58, 1 (Jan. 2012), 154–161. https://doi.org/10/dkkxwr

[5] Scott Carter, Jennifer Mankoff, Scott R. Klemmer, and Tara Matthews. 2008. Exiting the Cleanroom: On Ecological Validity and Ubiquitous Computing. *Human–Computer Interaction* 23, 1 (Feb. 2008), 47–99. https://doi.org/10.1080/07370020701851086 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07370020701851086.

[6] Blender Online Community. 2021. Blender - a 3D modelling and rendering package. http://www.blender.org

[7] Barrett Ens, Joel Lanir, Anthony Tang, Scott Bateman, Gun Lee, Thammathip Piumsomboon, and Mark Billinghurst. 2019. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies* 131 (Nov. 2019), 81–98. https://doi.org/10/ggrfdf

[8] Sebastian Friston, Ben J. Congdon, David Swapp, Lisa Izzouzi, Klara Brandstätter, Daniel Archer, Otto Olkkonen, Felix J. Thiel, and Anthony Steed. 2021. Ubiq: A System to Build Flexible Social Virtual Reality Experiences. (2021). https://doi.org/10/gn4vd5 MAG ID: 3217383357.

[9] Jerônimo Gustavo Grandi, Henrique Galvan Debarba, and Anderson Maciel. 2019. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 127–135. https://doi.org/10/gj27xd ISSN: 2642-5254.

[10] Carl Gutwin and Saul Greenberg. 2002. A Descriptive Framework of Workspace Awareness for Real-Time Groupware. *Computer Supported Cooperative Work (CSCW)* 11, 3 (Sept. 2002), 411–446. https://doi.org/10/c43wsg

[11] Joseph F. Hair, William C. Black, Barry J. Babin, and Rolph E. Anderson. 2009. *Multivariate Data Analysis* (7th edition ed.). Pearson, Upper Saddle River, NJ.

[12] Jon Hindmarsh, Mike Fraser, Christian Heath, Steve Benford, and Chris Greenhalgh. 2000. Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer-Human Interaction* 7, 4 (Dec. 2000), 477–509. https://doi.org/10/bbgzqx

[13] Ellen A. Isaacs and John C. Tang. 1994. What video can and cannot do for collaboration: A case study. *Multimedia Systems* 2, 2 (Aug. 1994), 63–73. https://doi.org/10.1007/BF01274181

[14] Brennan Jones, Yaying Zhang, Priscilla N. Y. Wong, and Sean Rintel. 2021. Belonging There: VROOM-ing into the Uncanny Valley of XR Telepresence. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (April 2021), 1–31. https://doi.org/10.1145/3449133

[15] Timothy A. Judge, Joyce E. Bono, Remus Ilies, and Megan W. Gerhardt. 2002. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology* 87, 4 (2002), 765–780. https://doi.org/10/bhfk7d Place: US Publisher: American Psychological Association.

[16] Kiyoshi Kiyokawa, Haruo Takemura, and Naokazu Yokoya. 1999. A collaboration support technique by integrating a shared virtual reality and a shared augmented reality. In *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028)*, Vol. 6. IEEE, 48–53. https://doi.org/10.1109/ICSMC.1999.816444

[17] Laan Labs. 2021. 3D Scanner App. https://apps.apple.com/us/app/3d-scanner-app/id1419913995

[18] Mark R. Leary and Robin M. Kowalski. 1993. The Interaction Anxiousness Scale: Construct and Criterion-Related Validity. *Journal of Personality Assessment* 61, 1 (Aug. 1993), 136–146. https://doi.org/10/b6kgfh Publisher: Routledge _eprint: https://doi.org/10.1207/s15327752jpa6101_10.

[19] Gun A. Lee, Seungwon Kim, Youngho Lee, Arindam Dey, Thammathip Piumsomboon, Mitchell Norman, and Mark Billinghurst. 2017. Improving Collaboration in Augmented Video Conference using Mutually Shared Gaze. *ICAT-EGVE 2017 - International Conference on Artificial Reality and Telexistence and Eurographics Symposium on Virtual Environments* (2017), 8 pages. https://doi.org/10/gm8b9d Artwork Size: 8 pages ISBN: 9783038680383 Publisher: The Eurographics Association.

[20] Gun A. Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. 2017. Mixed reality collaboration through sharing a live panorama. In *SIGGRAPH Asia 2017 Mobile Graphics & Interactive Applications (SA '17)*. Association for Computing Machinery, New York, NY, USA, 1–4. https://doi.org/10/ghk9ws

[21] Yuan Li, Feiyu Lu, Wallace S Lages, and Doug Bowman. 2019. Gaze Direction Visualization Techniques for Collaborative Wide-Area Model-Free Augmented Reality. In *Symposium on Spatial User Interaction (SUI '19)*. Association for Computing Machinery, New York, NY, USA, 1–11. https://doi.org/10.1145/3357251.3357583

[22] Paul Luff, Christian Heath, Hideaki Kuzuoka, Jon Hindmarsh, Keiichi Yamazaki, and Shinya Oyama. 2003. Fractured Ecologies: Creating Environments for Collaboration. *Human–Computer Interaction* 18, 1-2 (June 2003), 51–84. https://doi.org/10.1207/S15327051HCI1812_3 Publisher: Taylor & Francis _eprint: https://doi.org/10.1207/S15327051HCI1812_3.

[23] Jean-Luc Lugrin, Maximilian Landeck, and Marc Erich Latoschik. 2015. Avatar embodiment realism and virtual fitness training. In *2015 IEEE Virtual Reality (VR)*. 225–226. https://doi.org/10.1109/VR.2015.7223377 ISSN: 2375-5334.

[24] Mark McGill, Jan Gugenheimer, and Euan Freeman. 2020. A Quest for Co-Located Mixed Reality: Aligning and Assessing SLAM Tracking for Same-Space Multi-User Experiences. In *26th ACM Symposium on Virtual Reality Software and Technology*. ACM, Virtual Event Canada, 1–10. https://doi.org/10/ghwfq2

[25] Meta. 2021. Horizon Workrooms. https://www.oculus.com/workrooms

[26] Microsoft. 2022. Mixed Reality Toolkit (MRTK). https://github.com/microsoft/MixedRealityToolkit original-date: 2015-11-11T20:37:59Z.

[27] Jens Müller, Roman Rädle, and Harald Reiterer. 2017. Remote Collaboration With Mixed Reality Displays: How Shared Virtual Landmarks Facilitate Spatial Referencing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 6481–6486. https://doi.org/10.1145/3025453.3025717

[28] Shohei Nagai, Shunichi Kasahara, and Jun Rekimoto. 2015. LiveSphere: Sharing the Surrounding Visual Environment for Immersive Experience in Remote Collaboration. In *Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '15)*. Association for Computing Machinery, New York, NY, USA, 113–116. https://doi.org/10/gm8b9j

[29] Barbara J. Napier and Gerald R. Ferris. 1993. Distance in organizations. *Human Resource Management Review* 3, 4 (Dec. 1993), 321–357. https://doi.org/10.1016/1053-4822(93)90004-N

[30] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. 2016. Holoportation: Virtual 3D Teleportation in Real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST '16)*. Association for Computing Machinery, New York, NY, USA, 741–754. https://doi.org/10/gf2fpj

[31] Ye Pan, David Sinclair, and Kenny Mitchell. 2018. Empowerment and embodiment for collaborative mixed reality systems. *Computer Animation and Virtual Worlds* 29, 3-4 (2018), e1838. https://doi.org/10/gm8b8n _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cav.1838.

[32] Thammathip Piumsomboon, Arindam Dey, Barrett Ens, Gun Lee, and Mark Billinghurst. 2019. The Effects of Sharing Awareness Cues in Collaborative Mixed Reality. *Frontiers in Robotics and AI* 6 (2019). https://www.frontiersin.org/article/10.3389/frobt.2019.00005

[33] 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 *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173620

[34] Thammathip Piumsomboon, Gun A Lee, Andrew Irlitti, Barrett Ens, Bruce H Thomas, and Mark Billinghurst. 2019. On the shoulder of the giant: A multi-scale mixed reality collaboration with 360 video sharing and tangible interaction. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–17. https://doi.org/10.1145/3290605.3300458

[35] Thammathip Piumsomboon, Youngho Lee, Gun Lee, and Mark Billinghurst. 2017. CoVAR: a collaborative virtual and augmented reality system for remote collaboration. In *SIGGRAPH Asia 2017 Emerging Technologies (SA '17)*. Association for Computing Machinery, New York, NY, USA, 1–2. https://doi.org/10/gf2bqk

[36] Ronald Poelman, Oytun Akman, Stephan Lukosch, and Pieter Jonker. 2012. As if being there: mediated reality for crime scene investigation. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 1267–1276. https://doi.org/10/gm8b9q

[37] Beatrice Rammstedt and Oliver P. John. 2007. Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality* 41, 1 (Feb. 2007), 203–212. https://doi.org/10/djzd32

[38] Holger Regenbrecht and Thomas Schubert. 2021. Measuring Presence in Augmented Reality Environments: Design and a First Test of a Questionnaire. https://doi.org/10.48550/arXiv.2103.02831 arXiv:2103.02831 [cs].

[39] Danny Schott, Patrick Saalfeld, Gerd Schmidt, Fabian Joeres, Christian Boedecker, Florentine Huettl, Hauke Lang, Tobias Huber, Bernhard Preim, and Christian Hansen. 2021. A VR/AR Environment for Multi-User Liver Anatomy Education. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, Lisboa, Portugal, 296–305. https://doi.org/10/gk6nc9

[40] Mickael Sereno, Xiyao Wang, Lonni Besancon, Michael J Mcguffin, and Tobias Isenberg. 2020. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* (2020), 1–1. https://doi.org/10/gjkq7w Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[41] Khurram Shahzad, Usman Raja, and Syed Danial Hashmi. 2020. Impact of Big Five personality traits on authentic leadership. *Leadership & Organization Development Journal* 42, 2 (Jan. 2020), 208–218. https://doi.org/10/hfmf Publisher: Emerald Publishing Limited.

[42] Mel Slater, Domna Banakou, Alejandro Beacco, Jaime Gallego, Francisco Macia-Varela, and Ramon Oliva. 2022. A Separate Reality: An Update on Place Illusion and Plausibility in Virtual Reality. *Frontiers in Virtual Reality* 3 (2022). https://www.frontiersin.org/article/10.3389/frvir.2022.914392

[43] Mel Slater, Amela Sadagic, Martin Usoh, and Ralph Schroeder. 2000. Small-Group Behavior in a Virtual and Real Environment: A Comparative Study. *Presence: Teleoperators and Virtual Environments* 9, 1 (Feb. 2000), 37–51. https://doi.org/10.1162/105474600566600

[44] Mel Slater and Anthony Steed. 2000. A Virtual Presence Counter. *Presence* 9, 5 (Oct. 2000), 413–434. https://doi.org/10.1162/105474600566925 Conference Name: Presence.

[45] Rajinder S. Sodhi, Brett R. Jones, David Forsyth, Brian P. Bailey, and Giuliano Maciocci. 2013. BeThere: 3D mobile collaboration with spatial input. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Paris France, 179–188. https://doi.org/10.1145/2470654.2470679

[46] Anthony Steed, Lisa Izzouzi, Klara Brandstätter, Sebastian Friston, Benjamin James Congdon, Otto Olkkonen, Daniele Giunchi, Nels Numan, and David Swapp. 2022. Ubiq-Exp: A Toolkit to Build and Run Remote and Distributed Mixed Reality Experiments. *Frontiers in Virtual Reality* (2022). https://doi.org/10.3389/frvir.2022.912078 ISBN: 2673-4192 Publisher: Frontiers.

[47] Anthony Steed, Mel Slater, Amela Sadagic, Adrian Bullock, and Jolanda Tromp. 1999. Leadership and collaboration in shared virtual environments. In *Proceedings IEEE Virtual Reality (Cat. No. 99CB36316)*. IEEE, 112–115. https://doi.org/10.1109/VR.1999.756941

[48] Anthony Steed, William Steptoe, Wole Oyekoya, Fabrizio Pece, Tim Weyrich, Jan Kautz, Doron Friedman, Angelika Peer, Massimiliano Solazzi, Franco Tecchia, Massimo Bergamasco, and Mel Slater. 2012. Beaming: An Asymmetric Telepresence System. *IEEE Computer Graphics and Applications* 32, 6 (Nov. 2012), 10–17. https://doi.org/10.1109/MCG.2012.110 Conference Name: IEEE Computer Graphics and Applications.

[49] Anthony Tang, Omid Fakourfar, Carman Neustaedter, and Scott Bateman. 2017. Collaboration with 360° Videochat: Challenges and Opportunities. In *Proceedings of the 2017 Conference on Designing Interactive Systems*. ACM, Edinburgh United Kingdom, 1327–1339. https://doi.org/10.1145/3064663.3064707

[50] Theophilus Teo, Louise Lawrence, Gun A. Lee, Mark Billinghurst, and Matt Adcock. 2019. Mixed Reality Remote Collaboration Combining 360 Video and 3D Reconstruction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM, Glasgow Scotland Uk, 1–14. https://doi.org/10.1145/3290605.3300431

[51] Santawat Thanyadit, Parinya Punpongsanon, and Ting-Chuen Pong. 2019. ObserVAR: Visualization System for Observing Virtual Reality Users using Augmented Reality. In *2019 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 258–268. https://doi.org/10.1109/ISMAR.2019.00023 ISSN: 1554-7868.

[52] Lui Albæk Thomsen, Niels Christian Nilsson, Rolf Nordahl, and Boris Lohmann. 2019. Asymmetric collaboration in virtual reality: A taxonomy of asymmetric interfaces for collaborative immersive learning. *Tidsskriftet Læring og Medier (LOM)* 12, 20 (March 2019). https://doi.org/10.7146/lom.v12i20.109391 Number: 20.

[53] Alexander Toet, Tina Mioch, Simon N. B. Gunkel, Omar Niamut, and Jan B. F. van Erp. 2021. Assessment of Presence in Augmented and Mixed Reality: Presence in Augmented and Mixed Reality. In *CHI 2021: Workshop Evaluating User Experiences in Mixed Reality*. ACM SigCHI. https://doi.org/10.31234/osf.io/gvkwx

[54] Peng Wang, Xiaoliang Bai, Mark Billinghurst, Shusheng Zhang, Xiangyu Zhang, Shuxia Wang, Weiping He, Yuxiang Yan, and Hongyu Ji. 2021. AR/MR Remote Collaboration on Physical Tasks: A Review. *Robotics and Computer-Integrated Manufacturing* 72 (Dec. 2021), 102071. https://doi.org/10/gjwd7r

[55] Huahai Yang and Gary M. Olson. 2002. Exploring collaborative navigation: the effect of perspectives on group performance. In *Proceedings of the 4th international conference on Collaborative virtual environments (CVE '02)*. Association for Computing Machinery, New York, NY, USA, 135–142. https://doi.org/10.1145/571878.571899

[56] Kevin Yu, Gleb Gorbachev, Ulrich Eck, Frieder Pankratz, Nassir Navab, and Daniel Roth. 2021. Avatars for Teleconsultation: Effects of Avatar Embodiment Techniques on User Perception in 3D Asymmetric Telepresence. *IEEE Transactions on Visualization and Computer Graphics* (2021), 1–1. https://doi.org/10/gm8b8x Conference Name: IEEE Transactions on Visualization and Computer Graphics.

[57] Achim Zeileis, Francisco Cribari-Neto, Bettina Gruen, Ioannis Kosmidis, Alexandre Simas, and Andrea Rocha. 2021. betareg: Beta Regression. https://CRAN.R-project.org/package=betareg