Spectator View: Enabling Asymmetric Interaction between HMD Wearers and Spectators with a Large Display

FINN WELSFORD-ACKROYD, Computational Media Innovation Centre (CMIC), Victoria University of Wellington (VUW), New Zealand
ANDREW CHALMERS, CMIC, VUW, New Zealand
RAFAEL KUFFNER DOS ANJOS, University College of London, United Kingdom and CMIC, VUW, New Zealand
DANIEL MEDEIROS, University of Glasgow, United Kingdom and CMIC, VUW, New Zealand
HYEJIN KIM, CMIC, VUW, New Zealand
TAEHYUN RHEE, CMIC, VUW, New Zealand

Fig. 1. Example of the spectator view. Left: semi-immersed user in front of a large display looking around the virtual environment and interacting using motion controllers. Right: HMD wearer exploring the same environment, and communicating with their semi-immersed partner using pointer controls and their voice.

485

In this paper, we present a system that allows a user with a head-mounted display (HMD) to communicate and collaborate with spectators outside of the headset. We evaluate its impact on task performance, immersion, and collaborative interaction. Our solution targets scenarios like live presentations or multi-user collaborative systems, where it is not convenient to develop a VR multiplayer experience and supply each user (and spectator) with an HMD. The spectator views the virtual world on a large-scale tiled video wall and is given the ability to control the orientation of their own virtual camera. This allows spectators to stay focused on the immersed

Authors' addresses: Finn Welsford-Ackroyd, Computational Media Innovation Centre (CMIC), Victoria University of Wellington (VUW), 40 Taranaki Street, Wellington, New Zealand, finnwa24@gmail.com; Andrew Chalmers, CMIC, VUW, 40 Taranaki Street, Wellington, New Zealand, andrew.chalmers@vuw.ac.nz; Rafael Kuffner dos Anjos, University College of London, London, United Kingdom and CMIC, VUW, 40 Taranaki Street, Wellington, New Zealand; Daniel Medeiros, University of Glasgow, Glasgow, United Kingdom and CMIC, VUW, 40 Taranaki Street, Wellington, New Zealand; Hyejin Kim, CMIC, VUW, 40 Taranaki Street, Wellington, New Zealand; Taehyun Rhee, CMIC, VUW, 40 Taranaki Street, Wellington, New Zealand, taehyun.rhee@vuw.ac.nz.

2573-0142/2021/11-ART485 $15.00
https://doi.org/10.1145/3486951

user's point of view or freely look around the environment. To improve collaboration between users, we implemented a pointing system where a spectator can point at objects on the screen, which maps an indicator directly onto the objects in the virtual world. We conducted a user study to investigate the influence of rotational camera decoupling and pointing gestures in the context of HMD-immersed and non-immersed users utilizing a large-scale display. Our results indicate that camera decoupling and pointing positively impacts collaboration. A decoupled view is preferable in situations where both users need to indicate objects of interest in the scene, such as presentations and joint-task scenarios, as it requires a shared reference space. A coupled view, on the other hand, is preferable in synchronous interactions such as remote-assistant scenarios.

CCS Concepts: • **Human-centered computing → Displays and imagers**; **Pointing devices**; *Pointing*.

Additional Key Words and Phrases: large display; pointing; spectators; head-mounted display; asymmetric collaboration; motion filter

ACM Reference Format:
Finn Welsford-Ackroyd, Andrew Chalmers, Rafael Kuffner dos Anjos, Daniel Medeiros, Hyejin Kim, and Taehyun Rhee. 2021. Spectator View: Enabling Asymmetric Interaction between HMD Wearers and Spectators with a Large Display. *Proc. ACM Hum.-Comput. Interact.* 5, ISS, Article 485 (November 2021), 17 pages. https://doi.org/10.1145/3486951

1 INTRODUCTION

Head-mounted displays (HMDs) provide the wearer with a highly immersive virtual reality (VR) experience, allowing one to feel present in a Virtual Environment (VE). While a high sense of presence is key in completely virtual experiences, there are several use cases where one wants to be immersed but also communicate or interact with people in the same physical space (e.g., lectures, technical demonstrations, public presentations, and supervised training.) These are *asymmetric* use-case scenarios: the immersed user has a more active role, and non-immersed users can be considered *spectators* with limited affordances. In such cases, giving the spectators complete freedom is not desired. While allowing all users to have their own HMD is an option, these non-immersed users may have tasks that require them to be able to see both the virtual and real-world environments without wearing a headset. These include taking notes, capturing images, operating other equipment that is part of the experience, or engaging in communication with other non-immersed users. In asymmetric use-case scenarios, we must provide an asymmetric option for sharing the virtual experience.

Display mirroring does not give the spectator any control over what they view. Allowing the spectator to shift their focus away from what the VR user is looking at could be very useful for demonstrations and collaborative scenarios – they would be able to focus on features of the virtual environment, while still being co-located with the immersed VR user to facilitate communication. It restricts the spectator to verbal communication with the HMD user, having no way to visually communicate or point out objects of interest, which is highly detrimental in interactive scenarios between both parties. Display mirroring can also have undesirable effects on the spectator's experience: rapid head movements and rotation from the HMD wearer are not pleasing to watch and may induce simulator sickness [25]. Additionally, a regular display does not have any of the immersive qualities of an HMD. This makes it difficult for spectators to understand what the HMD wearer is experiencing since they are not themselves immersed – a common problem in the VR industry when attempting to communicate the VR experience to people not wearing an HMD.

In this paper, we evaluate the impact of interaction tools (pointing and camera rotational control) for a VR spectator using a large-scale tiled video wall. We propose a spectator system approach that combines these two, allowing a non-immersed person using the video wall to follow the experience of an immersed HMD user, as well as look around and communicate with them. We performed a user study evaluating the impact of each individual tool on both task performance, user preference, and other aspects of collaboration.

Previous works on collaboration between immersed and non-immersed users were designed for constrained scenarios such as specially designed games [14, 15, 38]. Our approach targets generic asymmetric collaboration/presentation scenarios where users have different affordances and tasks, but need to be able to share their experience, sense of immersion, and communicate about the environment (through the use of pointing gestures).

In summary, our contributions are: (1) A novel approach for asymmetric collaboration between an HMD wearer and spectators, (2) a user study that analyses the collaboration between spectators and HMD wearers, and (3) a system implementation of our approach covering display arrangement, spectator controllers, user interface, and a motion filter for the virtual camera.

Our results show that users preferred the approach combining both tools. It positively impacted the spectator's perceived immersion, improved communication between users, and allowed them to explore the environment independently.

2 RELATED WORK

This section addresses the relevant research on interacting with large-scale displays. We introduce the concept of large-scale displays, their immersive properties, and their role in VR technology. We then address the related work on interaction with VR content for these types of displays. Finally, we discuss collaboration, present the concept of asymmetrical collaboration and highlight the significant problems that this paper addresses.

2.1 Large-Scale Displays

Large-scale displays are used because they elicit a stronger sense of spatial presence than smaller screens [49] and are often compared with HMDs due to their immersive qualities [28]. They allow multiple co-located people to be immersed in VR together. Many different forms of large-scale displays have been developed, including highly immersive displays such as CAVEs [7] and large-scale domes [29]. Compared to HMDs, these screens offer users similar presence and spatial awareness levels [42] with reduced cybersickness side-effects [45].

There are many potential input methods for interacting with large-scale displays. Prior studies have examined the use of mobile phones [2], tablets [39], gloves [13], and gestures to select and manipulate 3D objects at a distance [32]. Input methods used alongside HMDs often employ a high level of motion tracking (up to six degrees of freedom (6DoF)) to capture the user's movement in as much detail as possible. This can increase the user's sense of presence in the VE [8]. Prior work have used motion controllers in non-HMD VR contexts to directly manipulate objects [4] and control the camera [51]. These studies also show that motion controllers need special consideration in implementation regarding the display type and the task being completed in the virtual environment.

In an HMD, the headset is used as the input device to control the camera [51]. There is a wide range of possible camera control methods for large-scale displays. Neng & Chambel [40] presented a dragging method for looking around in 360-videos, which has since been implemented in 360-video players on popular websites like Facebook and YouTube. Pavel et al. [43] presented a method for ensuring viewers can see important parts of a video. Users can press a button to be re-oriented in the direction of important content.

Previous works have investigated the use of pointing gestures as an interaction method [24], and have shown that different pointing techniques are effective at accomplishing different tasks. Pointing metaphors [51] differ in the method used to calculate the direction of the ray. In the image-plane technique, the virtual ray is cast from the user's eye and through their hand. In the laser metaphor, the 6DoF motion controller's position is used to calculate the intersection with the screen like a virtual laser pointer. Both metaphors are similar in efficiency, but the laser metaphor is better when dealing with multi-user scenarios, as the image-plane technique only looks correct to the person holding the controller [48]. For this reason, we used a similar metaphor to represent pointing gestures using a large-scale wall display.

2.2 Collaboration with Large-Scale Displays

Large-scale displays allow multiple people to interact and collaborate in a 3D space. Notable applications include design reviews in virtual environments [23] and virtual training scenarios [12]. Users can engage in a collaborative view of the VE, where multiple users can edit [52], comment, and collaborate. The large-scale display can present a first-person view of a single user [46] or an outside view, where participants can be seen at the same time.

The asymmetric nature of such systems has been recently explored in HMDs with immersed and non-immersed users [6, 16, 21, 34]. These systems use projections on the floor or tabletop [14, 15] to provide the non-immersed user with a view into the virtual world. These techniques improved the users' effectiveness at completing tasks and provided them with a sense of co-presence. Prior work has also enabled collaboration that converges users in VR and AR. A recent example is the Augmented Virtual Teleportation (AVT) system [44], where users can remotely collaborate using a 360° camera to link two separate spaces, enabling a remote VR user to collaborate with a local AR user. A coupled view has also been proven effective in remote assistance tasks between VR and AR users [41]. Asymmetric systems have also been configured with AR HMDs in collaboration with conventional displays [18].

Asymmetric collaboration is also possible between different display surfaces, such as the one described by Kunert et al. [26]. This system enables collaboration between a stereoscopic wall and a tabletop surface, which allows people to share the same reference space by using pointing gestures [26]. An important aspect of this is the users' need to be aware of each other's actions in order to enable effective communication and collaboration. ShareVR [15] tries to solve this problem with the use of projected floors and a screen mounted on a 6DoF controller, which enables the non-immersed user to glimpse into the shared 3D world of the HMD user. ReverseCave [22] enables spectators to observe the HMD user's experience. Alternatively, other recent work [6, 16] allows HMD users to collaborate by placing touch screens that are mounted onto the HMD.

However, none of these studies provide freedom of interaction as well as immersion to the non-HMD users. Collaboration with spectators has also been considered in augmented reality scenarios [19, 27], where users wearing AR headsets are able to share virtual information via tabletops, vertical displays, and point-of-view video streams to collaborators. It has been shown that sharing and exploring 3D data through different display types and tracking methods has a positive effect on task performance and co-presence. However, no prior work addresses the specific scenario of multiple co-located people using a large display alongside an HMD wearer. In this scenario, users may need to perform tasks which that require them to have an overview of elements in both real and virtual environments and allow for richer non-verbal communication, which is something that an HMD alone may not always be able to provide.

3 APPROACH AND SYSTEM DESIGN

As stated in Section 1, our work focuses on *asymmetric* use cases such as lectures, technical demonstrations, public presentations, and supervised training [18, 41]. In these situations the immersed HMD user has the main role of interacting and exploring the VE. The *spectator* has a different role and tasks, but must be able to: (1) share the sense of immersion in the VE, (2) communicate about the environment verbally and visually, and (3) loosely explore the environment around the HMD user to facilitate collaboration.

We designed an approach to work in this scenario. While some works that operate on asymmetric systems will offer equal affordances to both subjects [14, 15, 26, 44], we focus on the scenario of a shared experience where the focus is the HMD user.

A system implementing our approach can be seen in Figure 1. One immersed user (right) who utilizes a traditional HMD + controller kit is the leading participant of the experience, free to perform actions and move in the environment. The spectator (left) can passively visualize the VR experience in a large display, and can interact with the display using handheld controllers. The image displayed on the screen is subject to a camera filter to provide a better viewing experience when adapting the content to this alternative display media (detailed in Section 4).

3.1 Large screen display

While large screen displays have been bested in popularity by HMDs, they still serve as a viable alternative to provide a semi-immersive VR experience, eliciting a comparable sense of presence [42], but with increased sense of spatial awareness and reduced cybersickness side-effects [45] when filtered adequately (see Section 4.4). As one of our goals is to provide spectators with some of the immersion of the HMD user, we employ a large screen display to visualize the VE. Particularly, they are the best-suited alternative to showcase a VR experience to a **large audience**, as they provide the closest approximation possible to what the immersed person is experiencing, and allow all spectators to experience the virtual world in a similar way.

3.2 Rotational Camera Control

In order to grant the spectator control over their viewpoint, we added rotational camera control to the spectator system. We added a second camera which renders the spectator's view of the virtual world. By default, the spectator camera's position and orientation matches the primary HMD user's camera – so the spectator's screen displays what the HMD user sees in their headset. At any time, the spectator can take control of their camera to look in a different direction. This enhances the spectator's experience without affecting the HMD user's experience.

To enable spectators to control where the camera is looking, we used the "eyeball in hand" [51] technique. In this method, the spectator takes control of the virtual camera as if it were held in the user's hand. Using a controller to realize this metaphor allows the spectator to both naturally control the visualization, as well as easily switch into other tasks that may be necessary in the use case scenario (e.g. pointing, taking notes, communicating with other spectators, or supporting the HMD user), all while using a familiar device.

We considered allowing the spectator to have full 6DoF independent positional control of their camera. This would be more flexible for collaboration at the cost of increasing the system's complexity. It would also not be compatible with many first-person single-player experiences which are not designed with a second viewpoint in mind. As such, we decided to only allow for rotational control at this stage.

3.3 Pointing

Pointing was chosen as a method of visual interaction. It is a gesture often used in real life face-to-face interactions. In a virtual collaborative environment, it is an effective form to reduce the load on verbal communication [35] and focus more to the task at hand [20]. The design of the pointing tool was inspired by popular multiplayer games and are shown to be effective to indicate regions of interest in large-scale displays [24]. This use a laser approach where the 6DoF wand is tracked and is used to control position and rotation of the pointing ray.

To assist users in performing the task, we allowed them to place markers in the virtual scene. When the user presses the pointer button, a colored ray is cast in the direction they are pointing in the virtual world. When they let go of the button, a marker of the same color is placed at the location they last pointed at. The physical action of placing a marker using a motion controller is designed to mimic the real-world action of using a laser pointer as a presentation tool.

3.4 User Interface

To support natural communication between the users, we designed a *compass* UI element to help them quickly determine where their partner is looking. This is intended to eliminate the need to verbally communicate ordinal directions when discussing elements of the virtual environment. The compass takes the form of an arrow pointing in the direction that each user's partner is facing relative to their own position (see supplementary video). This is visible by both HMD user and spectator.

4 IMPLEMENTATION

In this section, we provide details of the implementation and setup of the spectator system.

4.1 Large-Scale Immersive Display

We configured the large-scale display to maximize its immersive properties. When configuring the video wall, the external field of view (FOV) is calculated with $FOV = 2 \times \arctan(x \div 2d)$, where x is the width or height of the display, and d is the viewer's distance from the display. The setup used in our user study (Fig. 1 left) had eight 55 inch displays arranged in a 4×2 grid. When d is 2000 mm, the horizontal FOV is 101,° and the vertical FOV is 38°. We then increased the horizontal fov by forming a curve with the displays, which has the benefits of increasing spatial presence [30] and improving the user experience by reducing viewing angles [9, 31].

4.2 Camera Control

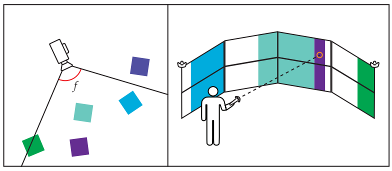
Two buttons are used for enabling and disabling camera control. One is the trigger, which gives the spectator control over their camera when held down. This resembles grabbing the camera to manipulate its orientation. The orientation of the controller when the trigger is pressed does not affect the camera's orientation. It is only the changes in controller position that are applied to the camera.

The other button toggles *free-cam*. If free-cam is disabled, the camera snaps back to following the HMD user's perspective when the spectator releases the trigger. If free-cam is enabled, the camera stays put. This feature was introduced so that the user could choose to fully separate their camera from the VR user without needing to hold the trigger. Additionally, it allows the user to perform "ratcheting" motions [51] to fully turn the camera around without maintaining an uncomfortable hand position. To ensure the spectator does not get confused, their heads-up display has an indicator that shows whether free-cam mode is enabled or disabled.

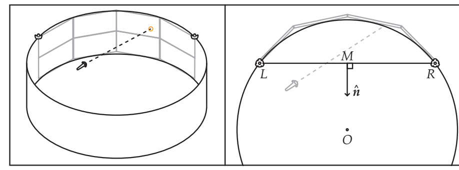
The spectator view utilizes the Unreal Engine 4 Virtual Reality Spectator Screen [11] feature to place a secondary camera into a VR scene. This camera renders to a render texture, which is shown on the immersive display as the spectator's view. The spectator camera control feature uses the button inputs to place the spectator camera into one of three states: locked to HMD, locked to spectator's motion controller, or stationary. The smoothing applied to the camera is achieved by storing a target orientation set by either the HMD or the motion controller. The camera's rotation is always interpolated towards the target orientation and never set directly, which has the effect of removing shakiness and smoothing out sharp changes in direction.

4.3 Pointing

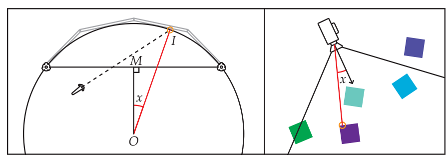
When the spectator points their controller at an object displayed on the screen and pulls the trigger, a virtual marker is placed on the object. If the user points the motion controller outside the display



(a) Left panel: Spectator's camera viewed from above. Right panel: The camera's view displayed on the screen as the spectator points at the purple column.



(b) Left panel: The video wall in the virtual world as a cylinder. Right panel: The cylinder viewed from above.



(c) Left panel: The controller's line trace collides with the cylinder. Right panel: Rotator x is added to the camera's forward vector to cast another line trace and place the marker.

Fig. 2. Details of the large-scale immersive display and interaction.

border, the marker is still placed at the corresponding position in the virtual world. This reduces the chance of unexpected behavior when the user aims for the screen and misses slightly. It even allows the user to point in the direction of objects that are not on-screen based on their mental model [36].

To achieve this behavior, we had to track the position of the large-scale display and map the spectator's view into the virtual world. The relationship between the 3D virtual scene and the 2D projection viewed by the spectator is depicted in Fig. 2(a). The left panel shows a top-down view of a hypothetical virtual scene. The spectator's camera is facing several tall square columns. The two lines extending from the camera represent its field of vision, where f is the FOV value in degrees. The right panel shows the camera's view of the scene as a 2D projection on the immersive display, with the user pointing their controller at the purple column.

To determine the position of the marker in the virtual world, the surface of the video wall display is tracked using Vive Trackers placed on the top corners of the display (shown as small crown-shaped objects in Fig. 2(a)). The positions of these trackers (along with the FOV value) are used to generate a large cylinder that fully surrounds the user. [1]

As shown in Fig. 2(b), we used the FOV value (f) and the positions of the left (L) and right (R) sides of the display to calculate the origin (O) and radius of the cylinder:

$$radius = \frac{|\frac{R-L}{2}|}{\sin(\frac{f}{2})} \quad and \quad O = M + (|O - M| \times \hat{n})$$

$$where \quad |O - M| = \frac{|\frac{R-L}{2}|}{\tan\frac{f}{2}} \quad and \quad \hat{n} = \frac{R - L}{|R - L|} \times [0, 0, -1]$$

To determine the position on the screen pointed to by the user, a single line trace [10] is used to simulate a laser emitted from the motion controller. To find the position to place the marker, another line trace must be cast from the spectator camera. This is because the spectator sees a 2D

[1] The same concept could also be applied to support dome displays if the cylinder was converted to a sphere.

projection of the virtual world from the perspective of the spectator camera. This perspective is not affected by the spectator's position in the real world. To hit the object pointed to by the spectator, the line trace is cast with an offset of x, which represents the angle between the forward vector and the impact point I (Fig. 2(c)).

4.4 Camera Filter

Camera movements such as rapid rotations [17] and shakiness [50] are more likely to cause simulator sickness. We applied a motion smoothing filter to the spectator's video feed to counteract these effects, similar to the one used by Tsubaki et al [50]. Rotations along the roll axis (when the VR user tilts their head to the side) are also hypothesized to induce motion sickness. This movement is not one that viewers are used to seeing displayed on screen: in both video games and film, camera rotations are usually restricted to tilting and yawing motions. As such, the camera filter removes rotations along the roll axis and keeps the spectator camera level.

5 USER STUDY DESIGN

Our implementation consists of the display setup, camera control, pointing, and the camera filter. We focus our study on the impact of camera control and pointing, as we assume an immersive display setup (with mitigated sickness through the camera filter) as the baseline which we aim to optimize against. Camera control and pointing are directly controlled by the spectator, which we evaluate.

We focused our study on rotational decoupling of the camera instead of a full decoupling (translation included) in order to evaluate the impact of an independent viewpoint, not also of an independent position. A full decoupling of the camera would introduce other variables to be evaluated (e.g., body representation, navigation metaphors), and while this might be useful in collaborative scenarios where both parties have equal autonomy, our evaluation focused on how much autonomy is desired in a spectator role.

The user test consisted of a task that required collaboration between an HMD user and a spectator. The setup used for the study can be seen in Fig. 3. We used a within-groups approach, where each pair completed a trial with four conditions. Each condition consisted of performing the task with one of the following combinations: (PC) both the pointer and camera decoupling enabled, (P) pointer only, (C) camera only, and the baseline (B) without pointer and camera decoupling. Furthermore, to maximize the amount of data collected from each pair, the four trials were repeated with the roles reversed. This means that each subject performed eight trials, acting as a spectator and a VR user in all evaluated conditions. The order of each of the four trials was performed using a balanced Latin squares arrangement to reduce the learning effect on the data.

5.1 Task Design

We designed a task that enabled an immersed (VR user) and a semi-immersed user (spectator) to interact and communicate to achieve a common goal, with the interaction aspects assigned to the VR user and a supportive role for the spectator. In this task, the users were placed in the center of a square room with a large grid on a wall and 40 unique moving cubes with various colors and patterns surrounding them (see Fig. 4(a)).

The VR user is able to pick up these cubes and place them in the grid (Fig. 4(b)). We provided the spectator with a printed screenshot of the grid with four cubes placed in it (Fig. 4(a)). The subjects' goal is to find the cubes displayed in the screenshot and place them in their corresponding locations. Participants had complementary goals where the VR user could interact with the virtual cubes, while the spectator would guide them to place the correct objects into the grid.

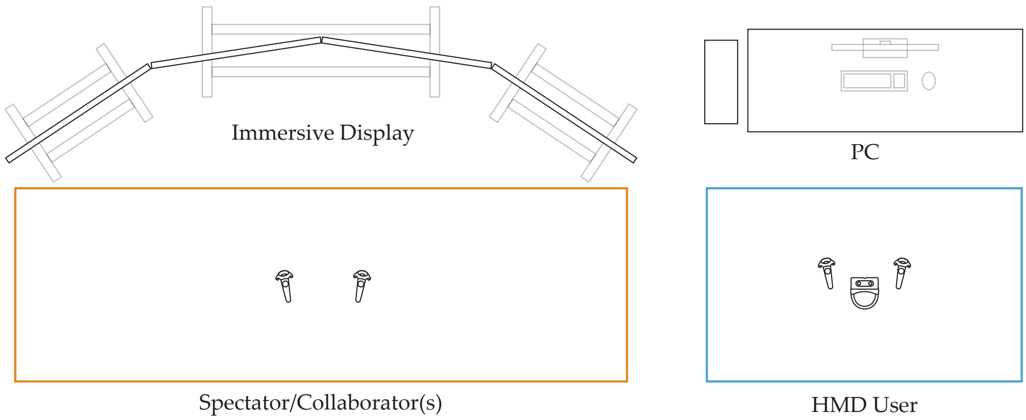


Fig. 3. Blueprint of the screen and room layout used for the user study.

To avoid users memorizing the room layout, the virtual cubes were randomly placed in the VE. Additionally, the cubes are randomly placed into the level at the beginning of each trial so that the users could not memorize the layout. The virtual cubes had complex texture patterns, making them difficult to describe verbally.

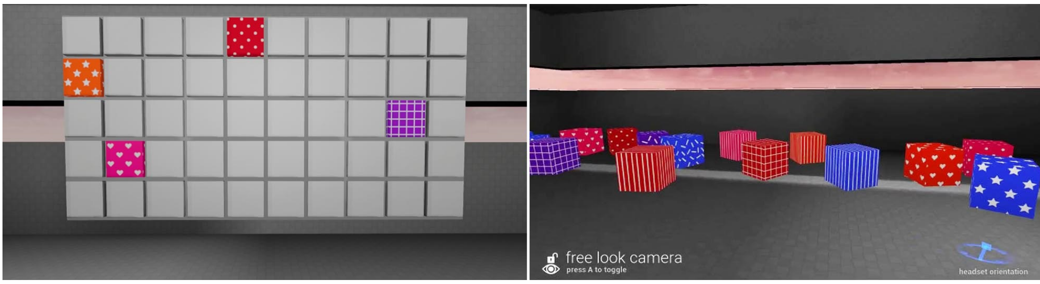
As discussed in previous work [50], using a camera filter reduced the effects of simulated sickness for spectators. Thus, we chose to keep the camera filter active for all the tested conditions. The hypothesis of the test was that the use of both the decoupled view and the pointing cues would make it easier for both subjects to complete the test compared to the baseline control condition.

5.2 Methodology

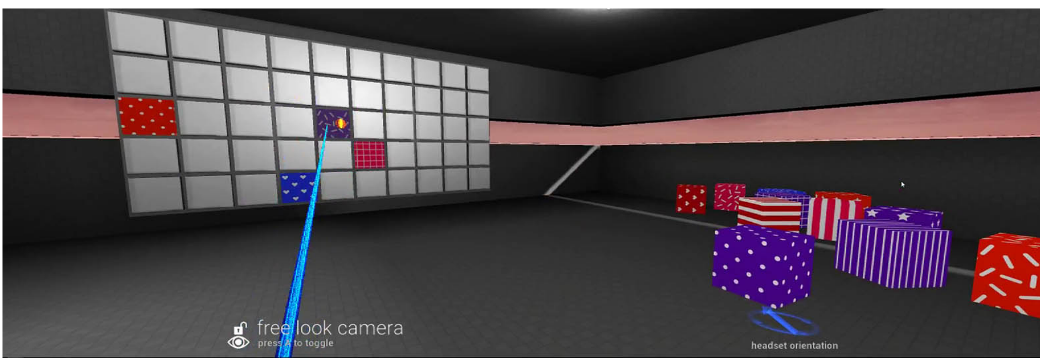
The study was conducted in a controlled environment where both users were located. A total of 24 participants completed the user study, where 75% were male, with ages ranging from 20-30, except for one that was more than 30 years old. Most of them (70%) had experience with VR systems. We greeted the participants and asked them to fill a pre-test questionnaire about their background and previous experience with VR systems. The user study was then explained to the participants. Each of them was randomly assigned the role of VR user or spectator and then asked to conduct the training to familiarize themselves with the system. Each session consisted of the four trials performed using a balanced Latin squares arrangement to avoid bias. The same four target grid layouts were used in each session (in randomized order). The target grid layout was then given to the spectator, and the time taken to complete the task was measured with a stopwatch. We also recorded how long each of the interactions took with the provided tools to confirm how much actual usage each of them had. After that, they were asked to fill a post-test questionnaire to gather subjective information about their user experience. The users then switched roles and conducted the same test with a different grid layout in all configurations.

5.3 Questionnaires

We used custom-made questionnaires that combined well-known metrics to evaluate elements such as usefulness of each tool, usability [5], the overall feeling of social presence [3] and presence [47]. The questionnaire was comprised of six 5-Likert scale questions, where 1 meant that the participant strongly disagreed with the statement and 5 that they strongly agreed with it. We focused on specific metrics of social presence in our user study: attentional allocation (AA) and perceived message



(a) Left: The test grid, as shown in one of the printed sheets given to the spectator. Right: The cubes in the room.



(b) The VR user places a cube into the grid. The orange sphere in the center of the grid is the spectator's marker.

Fig. 4. Virtual environment for the user study.

understanding (PMU) of the social presence questionnaire [3] to evaluate whether participants were able to focus on their actions (AA) and understand what they were doing (PMU). We also included questions related to Presence (i.e., the feeling of "being there" [47]) and overall task difficulty.

This questionnaire had the following statements: *1. It was easy to complete the task (tEasy), 2. I felt like I was actually there in the VE (Pres), 3. I remained focused on my partner throughout our interaction (AAMe), 4. My partner remained focused on me throughout our interaction (AAPart), 5. My thoughts were clear to my partner (PMUMe), 6. My partner's thoughts were clear to me (PMUPart).*

Then, we had questionnaires specific for each of the conditions, which included specific questions about the tools used in the condition and the role the user was experiencing, either VR or Spectator. Since our emphasis on the test was the usefulness of the system for the spectator, the questions used for the spectator's questionnaire asked about their partner's perceived effectiveness.

6 RESULTS

In this section, we present the results obtained from both quantitative and qualitative metrics for both users. Since the total test time is a continuous variable, we conducted a Shapiro-Wilk test to assess data normality. The samples were not normally distributed, so we applied the Friedman non-parametric tests to find statistical significance and Wilcoxon Signed Ranks Test post hoc tests for each pair of conditions, when statistically significant. The same tests were carried out for the questionnaires, given they used a discrete variable. In order to avoid Type II errors (false negatives), we did not use the Bonferroni correction on the pairwise tests, following the recommendation of previous research [1] that evaluated the impact of this type of correction on statistical analysis.

Moreover, in our case, we found that the Bonferroni correction would produce such errors where the Wilcoxon Signed Ranks tests with Bonferroni correction would contradict the results of the Friedman non-parametric test.

6.1 Subjective results

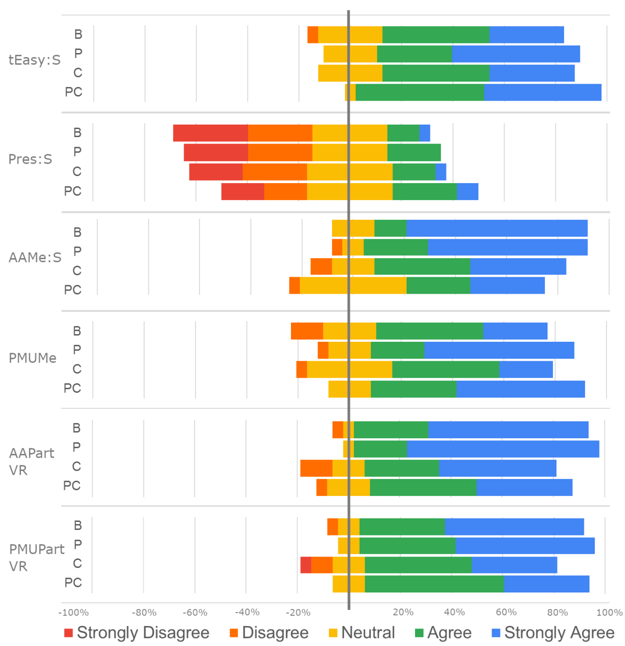


Fig. 5. Likert weighted graphs of specific questions in the user study.

We conducted data analysis to find statistical significance in the preferences questionnaires.

6.1.1 Social Presence. For social presence, we report the Attentional Allocation and Perceived Message Understanding (PMU) results.

Attentional Allocation: Regarding the spectator user, we found statistical significance in the question relating to attentional allocation ($\chi^2(3) = 16.6 p = 0.001$). When comparing between conditions, we found statistical significance in favor of those that did not have the decoupling of the camera. PC was both worse than P (Z=-3.176 p=0.001) and B (B Z=-2.982 p=-.003), and C was worse in relation to the P condition (Z=-2.140 p=0.032) and B (Z=-2.326 p=0.02). We highlight that the effect was worse with statistical significance when the camera was used in conjunction with the pointer.

For the VR User, we found statistical significance for the attentional allocation of the spectator (AAPart) ($\chi^2(3) = 12.140 p = 0.007$). By performing additional post hoc tests, we found a similar effect as the spectator user, where conditions that included the decoupling of the camera had worse results. PC was seen as worse than P (Z=-2.581 p=0.01) and B (Z=-2.310 p=0.021). We also found statistically significant results, with worse overall results for the C (camera-only) when compared to the P (pointer-only) condition (Fig. 5).

Perceived Message Understanding (PMU): We also found statistical significance in the PMU question, relating to the understanding of their actions by their partner (PMUPart) (($\chi^2(3) = 11.881 p =$

Table 1. Results obtained from the tools questionnaires. Results are presented as Median (Interquartile Range). PC indicates Pointing and controller, C Controller only and P Pointing only

VR User	PC	C	P	Spectator	PC	C	P
My partner's pointer tool was effective. (PC/P)	5(1)		5(1)	The pointer tool was effective. (PC/P)	5(1)		4(1)
It was useful being able to have my partner look in a separate direction. (PC/C)	5(2)	4(2)		The pointer tool was easy to use. (PC/P)	4(1)		4(2)
The compass tool displayed in my headset was effective. (PC/P/C)	4(2)	3.5(1.75)		The compass tool displayed on the screen was effective. (PC/P/C)	3(2)	3(2)	
The compass tool displayed in my headset was easy to use. (PC/P/C)	4(2)	4(2)		The compass tool was easy to use. (PC/P/C)	4(1.75)	4(2)	
				It was useful to be able to control my own camera. (PC/C)	5(1.75)	4(1.75)	
				It was easy to control my camera. (PC/C)	4(2)	4(1)	

0.008). In this question, users perceived they were clearer to their VR partner when using the pointer, as the PC was perceived as better in comparison with the B condition (Z=-2.001 p=0.045) and the P condition in comparison with both C (Z=-2.437 p=0.015) and B (Z=-2.195 p=0.028).

6.1.2 User Preferences. The tools were considered useful, except for the compass, where users had neutral responses. Users also pointed out that they found the decoupling of the camera useful on both the spectator and VR user roles, especially when paired with the pointer, but only with statistical significance in the VR role (Z=-3.000 p=0.003). The questions used for the spectator's questionnaire about their partner's perceived effectiveness are summarized in Table 1.

6.2 Task Performance

Regarding time, we did not find statistically significant results between each condition by performing pairwise comparisons.We did notice that on the conditions where spectators had the pointer tool, they were able to complete the task faster by an average of 13%. Our results also showed users put the available tools to use, averaging at 31% of the time (7% standard deviation) throughout all conditions. A summary of the results for each trial results can be seen in Fig. 6.

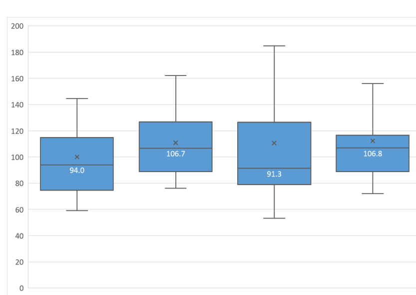


Fig. 6. Boxplots representing the total time per trial in seconds.

7 DISCUSSION

We found that the camera decoupling decreased the attentional allocation for the spectator (AAMe) and the perceived AA of the VR partner (AAPart) as noted by the lower results that include camera decoupling (PC and C). This informs us that on a task where both users need to collaborate over the same subject, independent rotational control of one's viewpoint is not desired. On the other hand, for a task such as the one proposed in our user study, decoupling views were used to "divide and conquer" when searching for each cube. This can be seen in the user study results. People noted the usefulness of camera decoupling in both roles, with a superior score (and statistically significant for the VR user) when paired with the pointer tool. This is due to the pointer tool being used to help re-sync views more quickly, which increased the perceived usefulness of camera decoupling. This is also explained by the relatively neutral scores of the compass's usefulness, showing that it did not assist the users in re-aligning their viewpoints as much as the pointer.

The pointer was shown to help each user understand each other's actions (PMU), which was found to have statistical significance. Overall, this tool was seen as the most useful by all users in the conditions it was available, given communication was a key part of successfully completing the task.

The condition with both camera decoupling and pointer (PC) also indicated to have a positive impact on users' presence. This condition was seen as easier by the users, especially when compared with the baseline condition. While the pointer tool played a bigger role in allowing users to achieve success, having the ability to control camera rotation at the same time had a positive impact on the users' experience, even if it did not translate into better performance.

We found no statistical significance in the difference of completion time for each condition, but we argue the importance of including additional observations collected during the experiments that indicate some tendencies to be explored in future work. Although users liked to have control over the camera, we can notice a slightly higher total time in the execution of the tasks (mitigated when the pointer was available) due to lost time re-aligning the viewpoints. Different strategies to illustrate the difference in viewpoints need to be investigated in future work, which would be more relevant in a scenario of full camera decoupling, where both users are equally active in the virtual environment, moving away from the "spectator" role. However, we believe the significant results in the usefulness of the camera decoupling tool showed that the additional time spent realigning the views was preferred over additional time looking for the objects (arguably a more demanding task).

In summary, we are able to say that each of the proposed tools can be used for specific scenarios. In the coupled view, users are able to better communicate synchronously (as seen by the PMU metric), so its use can be more effective in remote-assistant situations, as they can better understand their relationship and synchronize their view to perform a task more efficiently. The conjunction of a decoupled view and pointers can be used in scenarios such as presentations, when the spectator has to communicate with people that are also non-immersed in the virtual environment, and focus on details that the immersed user is not aware of. This can also be useful in a joint task-solving scenario, such as the one proposed by our user study. It is important to note that camera decoupling is an optional tool, and when not necessary, a coupled view with camera stabilization is used. This allows each of the users to interact concurrently and focus on different parts of the virtual scene.

8 CONCLUSION

We presented an approach for a spectator system in an asymmetric VR scenario where one user is immersed in the environment through an HMD, and the spectator is semi-immersed by interacting with a large-screen display. Interaction for the spectator is done through motion controllers where they can freely look around the virtual environment, decoupled from the HMD user's viewpoint. When the viewpoint is shared, the camera is stabilized with a filter. The spectator can also point at virtual objects through the large-screen display to communicate with the HMD user. We described in detail an implementation of the proposed approach and validated it with an exploratory user study.

Our results show that the proposed approach improved communication between both users, allowing each to focus on different areas of the environment to solve a collaborative task. The tools provided contributed positively to the spectator's presence in the system and were preferred by the test subjects. We achieved the asymmetric collaboration between HMD users and spectators through careful design considerations of user control, hardware setup, modes of communication, and the impact of presence. We also highlight the importance of a shared reference space for pointing gestures and verbal communication to work effectively. Additionally, we identified a list of scenarios for the use of each of the proposed collaboration strategies, depending on the level of collaboration needed. Future work that uses or gets inspiration from our system should also consider these aspects, and in turn, modify each component for effective asymmetric collaboration between spectators using a large-scale display and immersed users.

In future work, we will further explore the user interface provided by both users, increasing their awareness of where their partner is looking and their relative position. In our study, we only decoupled the rotation. It is possible to enable the spectator to translate away from the VR user's position. This would impact how users could observe one another from different positions. Future work will explore the impact of different types of display and motion tracking on the spectator. Exploring avatar representations of the spectator [37] and different tasks [33] inside the virtual environment may allow for richer communication between users.

ACKNOWLEDGMENTS

This project was supported by the Entrepreneurial University Programme funded by TEC and in part by the Smart Ideas project funded by MBIE in New Zealand. We also thank the anonymous reviewers for their insightful feedback.

REFERENCES

[1] Richard A. Armstrong. 2014. When to use the Bonferroni correction. *Ophthalmic and Physiological Optics* 34, 5 (2014), 502–508. https://doi.org/10.1111/opo.12131

[2] Teo Babic, Harald Reiterer, and Michael Haller. 2018. Pocket6: A 6DoF Controller Based On A Simple Smartphone Application. In *Proceedings of the Symposium on Spatial User Interaction* (Berlin, Germany) *(SUI '18)*. ACM, New York, NY, USA, 2–10. https://doi.org/10.1145/3267782.3267785

[3] Frank Biocca, C. Harms, and Jennifer Gregg. 2001. *The Networked Minds Measure of Social Presence: Pilot Test of the Factor Structure and Concurrent Validity*. Technical Report. Michigan State University. https://www.researchgate.net/publication/200772411_The_Networked_Minds_Measure_of_Social_Presence_Pilot_Test_of_the_Factor_Structure_and_Concurrent_Validity

[4] Doug A. Bowman, Brian Badillo, and Dhruv Manek. 2007. Evaluating the Need for Display-Specific and Device-Specific 3D Interaction Techniques. In *Virtual Reality (Lecture Notes in Computer Science)*, Randall Shumaker (Ed.). Springer, Berlin, Heidelberg, 195–204. https://doi.org/10.1007/978-3-540-73335-5_22

[5] John Brooke. 1996. *SUS: A 'Quick and Dirty' Usability Scale*. CRC Press, London, England. https://doi.org/10.1201/9781498710411-35

[6] Liwei Chan and Kouta Minamizawa. 2017. FrontFace: Facilitating Communication between HMD Users and Outsiders Using Front-Facing-Screen HMDs. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) *(MobileHCI '17)*. ACM, New York, NY, USA, 1–5. https://doi.org/10.1145/3098279.3098548

[7] Carolina Cruz-Neira, Daniel J. Sandin, Thomas A. DeFanti, Robert V. Kenyon, and John C. Hart. 1992. The CAVE: Audio Visual Experience Automatic Virtual Environment. *Commun. ACM* 35, 6 (June 1992), 64–72. https://doi.org/10.1145/129888.129892

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

[9] Alex Endert, Lauren Bradel, Jessica Zeitz, Christopher Andrews, and Chris North. 2012. Designing Large High-resolution Display Workspaces. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (Capri Island, Italy) *(AVI '12)*. ACM, New York, NY, USA, 58–65. https://doi.org/10.1145/2254556.2254570

[10] Epic Games. 2020. *UE4 Documentation: Using a Single Line Trace (Raycast) by Channel*. Epic Games. https://docs.unrealengine.com/en-US/InteractiveExperiences/Tracing/HowTo/SingleLineTraceByChannel/index.html

[11] Epic Games. 2020. *UE4 Documentation: Virtual Reality Spectator Screen*. Epic Games. https://docs.unrealengine.com/en-US/Platforms/VR/DevelopVR/VRSpectatorScreen/index.html

[12] Syafizwan Faroque, Ben Horan, Michael Mortimer, and Mulyoto Pangestu. 2016. Large-scale Virtual Reality micro-robotic cell injection training. In *2016 World Automation Congress (WAC)* (Rio Grande, PR, USA). IEEE, New York, NY, USA, 1–6. https://doi.org/10.1109/WAC.2016.7583006

[13] Paulo Gallotti, Alberto Raposo, and Luciano Soares. 2011. v-Glove: A 3D Virtual Touch Interface. In *Proceedings of the 2011 XIII Symposium on Virtual Reality* (Washington, DC, USA) *(SVR '11)*. IEEE, New York, NY, USA, 242–251. https://doi.org/10.1109/svr.2011.21

[14] 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)* (Osaka, Japan). IEEE, New York, NY, USA, 127–135. https://doi.org/10.1109/VR.2019.8798080 ISSN: 2642-5254.

[15] Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2017. ShareVR: Enabling Co-Located Experiences for Virtual Reality Between HMD and non-HMD Users. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. ACM, New York, NY, USA, 4021–4033. https://doi.org/10.1145/3025453.3025683

[16] Jan Gugenheimer, Evgeny Stemasov, Harpreet Sareen, and Enrico Rukzio. 2018. FaceDisplay: Towards Asymmetric Multi-User Interaction for Nomadic Virtual Reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. ACM, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3173628

[17] Jože Guna, Gregor Geršak, Iztok Humar, Jeungeun Song, Janko Drnovšek, and Matevž Pogačnik. 2019. Influence of video content type on users' virtual reality sickness perception and physiological response. *Future Generation Computer Systems* 91 (2019), 263 – 276. https://doi.org/10.1016/j.future.2018.08.049

[18] Kunal Gupta, Gun A Lee, and Mark Billinghurst. 2016. Do you see what I see? The effect of gaze tracking on task space remote collaboration. *IEEE transactions on visualization and computer graphics* 22, 11 (2016), 2413–2422.

[19] Kunal Gupta, Gun A. Lee, and Mark Billinghurst. 2016. Do You See What I See? The Effect of Gaze Tracking on Task Space Remote Collaboration. *IEEE Transactions on Visualization and Computer Graphics* 22, 11 (2016), 2413–2422. https://doi.org/10.1109/TVCG.2016.2593778

[20] Carl Gutwin, Saul Greenberg, and Mark Roseman. 1996. Workspace awareness in real-time distributed groupware: Framework, widgets, and evaluation. In *People and Computers XI*. Springer, London, 281–298.

[21] Hikaru Ibayashi, Yuta Sugiura, Daisuke Sakamoto, Natsuki Miyata, Mitsunori Tada, Takashi Okuma, Takeshi Kurata, Masaaki Mochimaru, and Takeo Igarashi. 2015. Dollhouse VR: a multi-view, multi-user collaborative design workspace with VR technology. In *SIGGRAPH Asia 2015 Emerging Technologies* (Kobe, Japan) *(SA '15)*. ACM, New York, NY, USA, 1–2. https://doi.org/10.1145/2818466.2818480

[22] Akira Ishii, Ippei Suzuki, Masaya Tsuruta, Shuta Nakamae, Junichi Suzuki, and Yoichi Ochiai. 2018. ReverseCAVE: CAVE-Based Visualization Methods of Public VR Towards Shareable VR Experience. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI EA '18)*. Association for Computing Machinery, New York, NY, USA, 1. https://doi.org/10.1145/3170427.3186593

[23] Ricardo Jota, Bruno R. de Araújo, Luís C. Bruno, João M. Pereira, and Joaquim A. Jorge. 2010. IMMIView: a multi-user solution for design review in real-time. *Journal of Real-Time Image Processing* 5, 2 (June 2010), 91–107. https://doi.org/10.1007/s11554-009-0141-1

[24] Ricardo Jota, Miguel A. Nacenta, Joaquim A. Jorge, Sheelagh Carpendale, and Saul Greenberg. 2010. A comparison of ray pointing techniques for very large displays. In *Proceedings of Graphics Interface 2010 (GI '10)*. Canadian Information Processing Society, CAN, 269–276. https://dl.acm.org/doi/10.5555/1839214.1839261

[25] Behrang Keshavarz and Heiko Hecht. 2011. Axis rotation and visually induced motion sickness: the role of combined roll, pitch, and yaw motion. *Aviation, space, and environmental medicine* 82, 11 (2011), 1023–1029.

[26] Andre Kunert, Tim Weissker, Bernd Froehlich, and Alexander Kulik. 2019. Multi-window 3d interaction for collaborative virtual reality. *IEEE transactions on visualization and computer graphics* 26, 11 (2019), 3271–3284.

[27] Andre Kunert, Tim Weissker, Bernd Froehlich, and Alexander Kulik. 2020. Multi-Window 3D Interaction for Collaborative Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics* 26, 11 (2020), 3271–3284. https://doi.org/10.1109/TVCG.2019.2914677

[28] Ed Lantz. 1996. The future of virtual reality: head mounted displays versus spatially immersive displays (panel). In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH '96)*. ACM, New York, NY, USA, 485–486. https://doi.org/10.1145/237170.237289

[29] Ed Lantz. 2007. A survey of large-scale immersive displays. In *Proceedings of the 2007 workshop on Emerging displays technologies: images and beyond: the future of displays and interacton* (San Diego, California, USA) *(EDT '07)*. ACM, New York, NY, USA, 1–es. https://doi.org/10.1145/1278240.1278241

[30] JJ-W Lin, Henry Been-Lirn Duh, Donald E. Parker, Habib Abi-Rached, and Thomas A. Furness. 2002. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceedings IEEE Virtual*

Reality 2002 (Orlando, FL, USA). IEEE, New York, NY, USA, 164–171. https://doi.org/10.1109/VR.2002.996519

[31] Lars Lischke, Sven Mayer, Katrin Wolf, Niels Henze, Harald Reiterer, and Albrecht Schmidt. 2016. Screen Arrangements and Interaction Areas for Large Display Work Places. In *Proceedings of the 5th ACM International Symposium on Pervasive Displays* (Oulu, Finland) *(PerDis '16)*. ACM, New York, NY, USA, 228–234. https://doi.org/10.1145/2914920.2915027

[32] Sreejith M, Siddharth Rakesh, Samik Gupta, Samprit Biswas, and Partha Pratim Das. 2015. Real-time hands-free immersive image navigation system using Microsoft Kinect 2.0 and Leap Motion Controller. In *2015 Fifth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG)* (Patna, India). IEEE, New York, NY, USA, 1–4. https://doi.org/10.1109/NCVPRIPG.2015.7489999

[33] Narges Mahyar and Melanie Tory. 2014. Supporting communication and coordination in collaborative sensemaking. *IEEE transactions on visualization and computer graphics* 20, 12 (2014), 1633–1642.

[34] Christian Mai, Lukas Rambold, and Mohamed Khamis. 2017. TransparentHMD: Revealing the HMD User's Face to Bystanders. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia* (Stuttgart, Germany) *(MUM '17)*. ACM, New York, NY, USA, 515–520. https://doi.org/10.1145/3152832.3157813

[35] Marilyn M. Mantei, Ronald M. Baecker, Abigail J. Sellen, William A. S. Buxton, Thomas Milligan, and Barry Wellman. 1991. Experiences in the Use of a Media Space. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New Orleans, Louisiana, USA) *(CHI '91)*. Association for Computing Machinery, New York, NY, USA, 203–208. https://doi.org/10.1145/108844.108888

[36] Anders Markussen, Sebastian Boring, Mikkel R. Jakobsen, and Kasper Hornbæk. 2016. Off-Limits: Interacting Beyond the Boundaries of Large Displays. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. ACM, New York, NY, USA, 5862–5873. https://doi.org/10.1145/2858036.2858083

[37] Daniel Medeiros, Rafael K. dos Anjos, Daniel Mendes, João Madeiras Pereira, Alberto Raposo, and Joaquim Jorge. 2018. Keep My Head on My Shoulders! Why Third-Person is Bad for Navigation in VR. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 16, 10 pages. https://doi.org/10.1145/3281505.3281511

[38] Daniel Medeiros, Eduardo Ribeiro, Peter Dam, Rodrigo Pinheiro, Thiago Motta, Manuel Loaiza, and Alberto B. Raposo. 2012. A Case Study on the Implementation of the 3C Collaboration Model in Virtual Environments. In *2012 14th Symposium on Virtual and Augmented Reality* (Rio de Janeiro, Brazil). IEEE, New York, NY, USA, 147–154. https://doi.org/10.1109/SVR.2012.28

[39] Daniel Medeiros, Lucas Teixeira, Felipe Carvalho, Ismael Santos, and Alberto Raposo. 2013. A tablet-based 3D interaction tool for virtual engineering environments. In *Proceedings of the 12th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry* (Hong Kong, Hong Kong) *(VRCAI '13)*. ACM, New York, NY, USA, 211–218. https://doi.org/10.1145/2534329.2534349

[40] Luís A. R. Neng and Teresa Chambel. 2010. Get around 360° Hypervideo. In *Proceedings of the 14th International Academic MindTrek Conference: Envisioning Future Media Environments* (Tampere, Finland) *(MindTrek '10)*. ACM, New York, NY, USA, 119–122. https://doi.org/10.1145/1930488.1930512

[41] Ohan Oda, Carmine Elvezio, Mengu Sukan, Steven Feiner, and Barbara Tversky. 2015. Virtual Replicas for Remote Assistance in Virtual and Augmented Reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 405–415. https://doi.org/10.1145/2807442.2807497

[42] Emilee Patrick, Dennis Cosgrove, Aleksandra Slavkovic, Jennifer A. Rode, Thom Verratti, and Greg Chiselko. 2000. Using a Large Projection Screen As an Alternative to Head-mounted Displays for Virtual Environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. ACM, New York, NY, USA, 478–485. https://doi.org/10.1145/332040.332479

[43] Amy Pavel, Björn Hartmann, and Maneesh Agrawala. 2017. Shot Orientation Controls for Interactive Cinematography with 360 Video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology* (Québec City, QC, Canada) *(UIST '17)*. ACM, New York, NY, USA, 289–297. https://doi.org/10.1145/3126594.3126636

[44] Taehyun Rhee, Stephen Thompson, Daniel Medeiros, Rafael Kuffner Dos Anjos, and Andrew Chalmers. 2020. Augmented Virtual Teleportation for High-Fidelity Telecollaboration. *IEEE Transactions on Visualization and Computer Graphics* 26, 5 (2020), 1923–1933. https://doi.org/10.1109/tvcg.2020.2973065

[45] Sarah Sharples, Sue Cobb, Amanda Moody, and John R. Wilson. 2008. Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems. *Displays* 29, 2 (2008), 58 – 69. https://doi.org/10.1016/j.displa.2007.09.005

[46] Andreas Simon. 2005. First-person Experience and Usability of Co-located Interaction in a Projection-based Virtual Environment. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* (Monterey, CA, USA) *(VRST '05)*. ACM, New York, NY, USA, 23–30. https://doi.org/10.1145/1101616.1101622

[47] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144.

[48] Maurício Sousa, Rafael Kuffner dos Anjos, Daniel Mendes, Mark Billinghurst, and Joaquim Jorge. 2019. Warping Deixis: Distorting Gestures to Enhance Collaboration *(CHI '19)*. ACM, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300838

[49] Desney S. Tan, Darren Gergle, Peter Scupelli, and Randy Pausch. 2003. With Similar Visual Angles, Larger Displays Improve Spatial Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. ACM, New York, NY, USA, 217–224. https://doi.org/10.1145/642611.642650

[50] Ikuko Tsubaki, Toshiya Morita, Takahiro Saito, and Kiyoharu Aizawa. 2005. An adaptive video stabilization method for reducing visually induced motion sickness. In *IEEE International Conference on Image Processing 2005* (Genova, Italy), Vol. 3. IEEE, New York, NY, USA, III–497. https://doi.org/10.1109/ICIP.2005.1530437

[51] Colin Ware and Steven Osborne. 1990. Exploration and Virtual Camera Control in Virtual Three Dimensional Environments. *SIGGRAPH Comput. Graph.* 24, 2 (Feb. 1990), 175–183. https://doi.org/10.1145/91394.91442

[52] Haijun Xia, Sebastian Herscher, Ken Perlin, and Daniel Wigdor. 2018. Spacetime: Enabling Fluid Individual and Collaborative Editing in Virtual Reality. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. ACM, New York, NY, USA, 853–866. https://doi.org/10.1145/3242587.3242597

Received February 2021; revised May 2021; accepted June 2021