Exploring Bi-Directional Pinpointing Techniques for Cross-Reality Collaboration

8 authors, including:

Daniel Simões Lopes
Instituto Superior Técnico
106 PUBLICATIONS   **512** CITATIONS

SEE PROFILE

Robbe Cools
KU Leuven
4 PUBLICATIONS   **6** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   IT-MEDEX - Interactive Tablets for Collaborative Scenarios Related to 3D Medical Image Exploration   View project

Project   Molecular Visualization and Simulation   View project

Exploring Bi-Directional Pinpointing Techniques for Cross-Reality Collaboration

Priyanka Pazhayedath*
Department of Computer
Science
KU Leuven

Pedro Belchior*
Instituto Superior Técnico
University of Lisbon

Rafael Prates*
Instituto Superior Técnico
University of Lisbon

Filipe Silveira*
Instituto Superior Técnico
University of Lisbon

Daniel Simões Lopes †
INESC ID
Instituto Superior Técnico
University of Lisbon

Robbe Cools‡
Department of Computer
Science
KU Leuven

Augusto Esteves §
ITI / LARSyS
Instituto Superior Técnico
University of Lisbon

Adalberto L. Simeone‡
Department of Computer
Science
KU Leuven

Figure 1: The three uni-directional Cross-Reality pinpointing techniques explored in the first study (from left to right): no visual feedback, audio only via *voice* clips; a *highlight* outlining the object (in white); and a three-dimensional arrow.
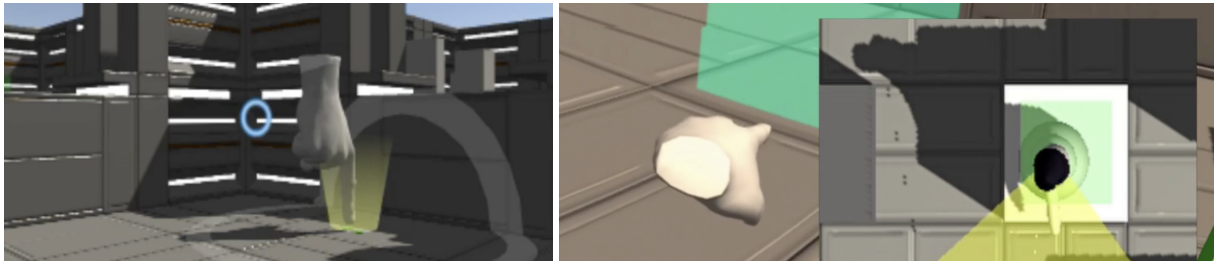

Figure 2: Two of techniques explored in the second study in the context of bi-directional Cross-Reality pinpointing. Left – a pointing action from the tablet user, from the perspective of the user in VR. Inspired by Ibayashi et al. [12], we have added a pulse element that travels the scene and enables vibrotactile feedback when passing through the VR user. This enables she or he to be aware of this action even if facing some other direction. Right – a pointing action by the VR user (tracked via the handheld controllers), and how this is represented to the tablet user.

ABSTRACT

Virtual Reality (VR) technology enables users to immerse themselves in artificial worlds. However, it isolates users from the outside world and impedes them from collaborating with other users who might be outside of the VR experience and vice-versa. We implemented two systems where we explore how such an external user in the real world can interact across realities with a user immersed in virtual reality, either locally or remotely, in order to to share pinpoint locations. In the first we investigate three cross-reality techniques for the external user to draw the attention of their VR counterpart on specific objects present in the virtual environment (Voice, Highlight, and Arrow). Participants performed better overall and preferred the Arrow technique, followed by the Highlight technique. In the second system we expand on these two techniques to explore an even

starker cross-reality interaction between users in VR and users interacting via a tablet computer to direct each other to pinpoint objects in the scene. We adapted the previous two techniques and implemented two others (Vision cone, Pointing) that support bi-directional communication between users. When it comes to bi-directional pinpointing, VR users still showed preference for the Arrow technique (now described as Pointing in Giant mode), while mobile users were split between the Vision cone and the Highlight techniques.

Index Terms: Human-centered computing—Collaborative and social computing—Empirical studies in collaborative and social computing; Human-centered computing—Human computer interaction (HCI)—Interaction techniques; Human-centered computing—Human computer interaction (HCI)—Interaction paradigms—Virtual reality

1 INTRODUCTION

In conventional Virtual Reality (VR) experiences, the user is almost completely immersed in an artificial world. In contexts where multiple users collaborate across different immersion levels, communication "across realities" becomes necessary. This is especially challenging between users that are immersed in VR and those who are "outside" VR. They each see a different reality and are not aware of what is happening in the other, e.g., to what they are seeing or

*e-mail: {priyankaprakasan, pedro.fls.belchior, rafaelprates1998, flip-slv98}@gmail.com

†e-mail: daniel.lopes@inesc-id.pt

‡e-mail: {robbe.cools, adalberto.simeone}@kuleuven.be

§e-mail: augusto.esteves@tecnico.ulisboa.pt

pointing at.

The increased affordability of "XR" (Virtual / Augmented / Mixed Reality) headsets has led these "Cross-Reality" (CR) collaboration scenarios to become more common. While VR applications aimed at immersed users do provide some degree of awareness of the outside world (e.g., SteamVR's *Chaperone* grid), support for interaction going in the opposite direction (awareness of a VR user's virtual environment and point of view; pinpointing points of interest) is still rudimentary. Without explicit support, external users must rely either on mirror screens (associated to the VR user's perspective or not) or explicit voice instructions (e.g., *"do you see an object that looks like a ...?"*) to better understand what the VR user is seeing or pointing towards.

To investigate this further, we first designed two unidirectional Cross-Reality pinpointing techniques that support external users (EUs) in collaborating with immersed user (IUs). These techniques use different modalities to focus the IU's attention on an object in the virtual environment (VE). We thus compared a disruptive technique, implemented as a three-dimensional *arrow*, that in the IU's reality points towards the object chosen by the EU, and a less overt *highlighting* technique to a baseline method consisting of synthesised voice-clips. These were then evaluated in a user study where the IU had to identify and select the object in the shortest time possible. Successively, we ran a follow-up study where the previous two Cross-Reality pinpointing techniques were expanded to support bidirectional interaction, in additional to a new *vision cone* technique. These were explored in the context of real-time puzzles where the EU had a bird's eye view over the VE, with the ultimate goal of minimising the need for speech.

Our results show a clear advantage for disruptive techniques in terms of task completion times when the EU is the originator of the interaction intent. In the other direction, results suggest that tablet EUs who receive the interaction intent from the IU showed appreciation for both disruptive techniques such as the vision cone and less disruptive techniques such as object highlighting.

2 RELATED WORK

Intra-Reality collaboration, that is, activities between multiple users in the same VE are typically referred to as examples of Collaborative Virtual Environments (CVEs) [4, 19]. In this case users are all immersed through the same technology-enabling method. Literature provides various examples of such CVEs. CollaVR [16] presents a way of collaboratively reviewing 360 video content. Spacetime [22] presents a CVE for VR users, and supports multi-user interactions with interaction concepts such as Container, Parallel Objects and Avatar Objects.

The increased availability of different methods to immerse one-self in a mixed reality environment has given rise to opportunities for "Cross-Reality" collaboration. With Cross-Reality (CR) we refer to scenarios in which the collaborators are at different points of the Reality-Virtuality continuum [15]. When users are no longer immersed in the same reality, but still retain the ability to interact with each other across different realities, we refer to these asymmetric interactions as examples of Cross-Reality Interactions. These can be differentiated on the basis of which "realities" are involved in the interaction.

Concerning AR and external users, Stafford et al. [20] investigated a system enabling external users working on a tabletop surface that captured the touch gestures performed on the tabletop and relayed them as virtual content overlaid on the display of AR users. The MagicBook by Billinghurst et al. [1] is a book capable of seamlessly functioning across different realities (reality, AR, and VR). Piumsomboon et al. compared different methods for sharing attention cues between AR and VR users, finding the combination of FoV frustum and head-gaze ray to be the most effective [17].

Examples of interactions involving VR and external users are

Figure 3: External UI with IU avatar (i.e., sphere) during the uni-directional study.

the following. FaceDisplay [8] allows external users to watch the virtual space through touch screen displays worn around the VR user's headset and interact with them either by touching the displays or via gestures. A similar approach of mounting a screen to the HMD is also presented in FrontFace [3]. ShareVR [7] uses a floor projection to enable external users to see into the VE. Additionally, a portable screen was also used to create a "window" into the VE. HMD Light [21] provides an alternative to a static projector by mounting a projector onto the HMD. The projector is motor-controlled and projects onto the floor in front of the VR user. A depth camera is also used to enable external users to interact with the projection. RealityCheck [11] explores projecting the VE into the real environment, but also investigates a method of compositing the real environment into the VE. This enables the immersed user to interact with external users and objects in the real environment. Interruptions and notifications from outside the VE are explored in NotifiVR [5]. These interruptions can be physical, such a person or a pet, or digital, such as a text message or voice call. TransceiVR [13] enables communication between an immersed VR user and an external tablet user. The latter can view the VE from the viewpoint of the immersed user and can then freeze this view to draw an annotation which is shown to the immersed user in the VE.

Our work adds to the existing body of CR literature the results of two empirical studies we conducted to compare different, yet well-known, techniques to focus the attention of the IU by the EU when performing selection/pinpointing tasks techniques. Specifically, we focused on scenarios supporting asymmetric collaboration between immersed and external users (using either a desktop application or a tablet).

3 UNI-DIRECTIONAL TECHNIQUES IMPLEMENTATION

To enable CR collaboration, we designed a networked application which presents an asymmetric user interface whose functionality depends on whether its user is immersed or not. Both users were represented as abstract spheres in each other's view (see Fig. 3-4). The environment chosen for the study scenario was a living room.

3.1 Desktop and VR interfaces

The user interface for the EU consists of a conventional 3D desktop application. It provides an independent viewpoint into the 3D scene. The EU can explore the scene via a conventional WASD and mouse interface. The user interface for the IU, instead, provides a first-person VR perspective into the same living room environment. The environment was made to fit into a room-scale area of $4\,m \times 4\,m$. The two prototype applications were developed using Unity (2018.3.11). The UNET framework was used to establish the connection between VR and desktop application. During the experiment, the IU acted as the server and the EU as the client.

We adapted two existing interaction techniques for CR and compared them with a technique based on voice communication, in-

Figure 4: VR UI with EU avatar (i.e., sphere) during the uni-directional study.

tended to represent the current baseline method.

Voice — In this interaction technique, the EU recorded an audio description of the object to be notified to the IU. A voice recognition algorithm converted it as a text message, which is then sent over the network. A voice synthesizer then played the audio back to the IU. This approach was chosen to make sure that the way these descriptions were communicated remained consistent between trials.

Highlight — The EU highlights the chosen object in their 3D environment via the mouse through simple ray-cast intersection. Data about the chosen object is then sent to the IU application, which causes the corresponding object to become highlighted.

Arrow — The EU selects an object in their 3D environment with the mouse, as before. However, upon receiving the message, a moving 3D Arrow is instantiated in front of the IU's headset. The Arrow will then move and rotate towards the corresponding object in the VE.

4 UNI-DIRECTIONAL USER STUDY

We conducted a within-subject user study to assess the performance and qualitative measures of the two CR interaction techniques compared to the baseline method consisting of voice clips. The results of this study and the following are discussed together in Section 7.

4.1 Participants and Apparatus

Eight participants took part in the experiment (3 females) as IUs. They were aged between 19 and 30. All of them were recruited via email and social media. The experimenter acting as the EU interacted via a 3D Desktop Application by using the mouse and keyboard, whereas IUs wore a HTC Vive connected to another machine. Both were part of the same wireless network.

4.2 Experimental Design

We had 8 participants, thus 24 trials were conducted (8×3 *techniques*) counter balanced using a Latin square. 3D models of 15 different books were evenly distributed and arranged in the shelves of living room of both External and VR user Interfaces. We measured task completion time and qualitative measures (workload, NASA-TLX [9]; user experience, UEQ [14]; simulator sickness, SSQ [6]; user subjective preferences via a non-validated questionnaire) as dependent variables.

4.3 Task

Participants acted as IUs. They were asked to identify a specific book. The experimenter, acting as the EU, initiated the search task via the 3D desktop application. They chose a random different book for each trial among those available in the living room VE, using the technique assigned to the particular trial. After each trial, while the 3D models of the books remained in the same position, their textures, and thus the book covers, were randomly shuffled.

Figure 5: Study setup for participants in the second study, exploring bi-directional pinpointing techniques between immersed (VR) and external users (interacting via a tablet computer).

4.4 Procedure

The experimenter explained participants the higher level goals of the task, consisting in evaluating different CR techniques. They then signed a consent form before participating.

Each trial began when the EU initiated the search task using the interaction technique assigned to the trial. Through each CR interaction technique, a particular book was chosen that the IU had to find. Visual feedback was present for the *Highlight* and *Arrow* techniques, and absent in the case of *Voice*. Once the book was found by the IU, they confirmed their selection by pointing at it with their controller and pressing the trigger button. After the conclusion of the experiment, the participant was asked to fill the previously indicated questionnaires. The average duration of the whole experiment per participant was about 30 minutes.

4.5 Results

We have evaluated the performance of the three considered CR techniques (Voice, Highlight and Arrow) using measures such as the task completion time (TCT) and other qualitative metrics such as user experience, workload, simulator sickness and a non-validated users custom subjective questionnaire for ranking the interaction techniques and rating the external user interactions.

Participants completed the task using the Arrow technique in a mean time of 3.97 s ($SD = 1.66$), Highlight with a mean time of 9 s ($SD = 9.05$) and Voice with a mean time of 17.23 s ($SD = 12.62$). Due to the non-normality of the data, we conducted a Kruskal-Wallis analysis on the performance data collected. Significant differences ($p < .01$) were found in terms of interaction technique between Voice vs Arrow and Voice vs Highlight ($p = .03$). Pair-wise comparisons between Arrow and Highlight were non-significant ($p = .06$).

We measured user experience using the UEQ. Arrow received a mean score of 1.90 ($SD = 0.56$), Voice 2.75 ($SD = 1.53$) and Highlight 2.96 ($SD = 1.53$). An ANOVA did not show any significant differences between the interaction techniques. We then analysed the workload associated with each technique through the NASA-TLX questionnaire. The Arrow technique received a mean overall workload score of 16.25 ($SD = 18.78$), Voice 28.38 ($SD = 20.84$) and Highlight 23.44 ($SD = 23.93$). However, these differences were not significant. In the Simulator Sickness Questionnaire, Voice received an average score of 0.12 ($SD = 0.10$), Highlight 0.08 ($SD = 0.09$) and Arrow 0.10 ($SD = 0.14$). There were no significant differences in terms of SSQ scores.
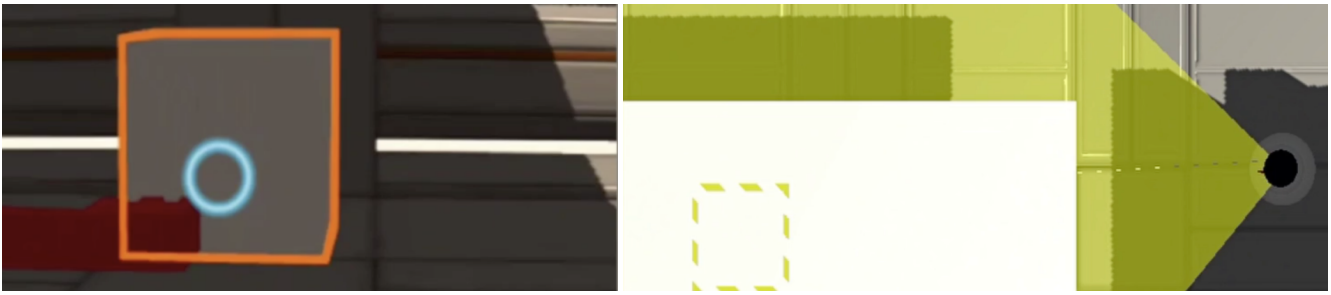
Figure 6: The highlighting technique from both the VR and tablet users' perspectives (left and right, respectively), in response to the VR user's head pointing (blue circle). The tablet user will perceive this even if occluded by some other element in the scene (right). Unrelated to this is a red key, which is not manipulable by the VR user in its current state (left).
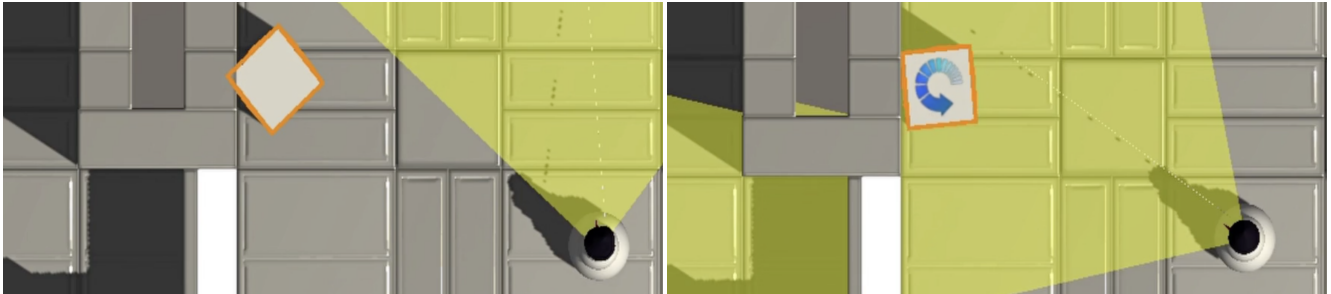

Figure 7: The view cone of the VR user as seen by the tablet counterpart. When this goes over an interactable element, a hint is often overlaid onto it (e.g., rotate the element – right).

5 BI-DIRECTIONAL TECHNIQUES IMPLEMENTATION

Following these results, and inspired by Billinghurst et al. [1], we expanded on these Highlight and Arrow techniques by exploring an even starker asymmetric task between IU and EU (Figure 5). These follow-up pinpointing techniques were designed as surrogates to voice communication – which is not feasible in many mobile or remote collaboration settings – and studied in the context of a series of puzzles where the real-time collaboration between pairs of users is critical (e.g., when some puzzle element is not perceivable by the tablet user, who sees the scene in either a top-down or isometric view following a Giant mode perspective [18]).

5.1 VR Users

Users interacting from VR were represented via a simple avatar that could communicate with tablet users through implicit and explicit techniques. Regarding the former, we re-implemented the successful **highlight** technique from the previous study that created a yellow-dashed contour around the objects being gazed at (via head-pointing) – this contour was visible even if these objectives were occluded from the tablet user's perspective (see Figure 6 – right). We also implemented a **vision cone** that showed the tablet user what was in viewing range of their VR counterpart. When this cone was over an object, some hints were provided to the tablet user on how to solve the puzzle (e.g., a rotate icon over an object – see Figure 7 – right). Finally, IU could also explicitly **point** to points-of-interest, tracked via the handheld controllers and represented by the avatar's arms (see Figure 2 – right). Taken together, these could be used to communicate which actions should be carried out by the EU: from dragging an object from A to B, to scaling or rotating it. Objects that were only interactable by the VR user were shown in green, adding an additional requirement for collaboration.

5.2 Tablet Users

Users interacting via the tablet computer had access to two main functions. First, they could alter the **viewing perspective** of the entire scene from top-down to isometric via a button on the top-left corner of the interface. Because they were in this giant mode, they had no representation in the scene other than **pointing** to a point-of-interest via a touch event. If this was done over a general location, and following the giant metaphor, then it was represented not by an arrow as in the first implementation but by a giant fist with the index finger extended. This was inspired by Ibayash et al. [12] and improved with a pulse animation of an ever-expanding ring that would produce a vibrotactile feedback when crossing the VR user's avatar (via the handheld controllers). Such technique was considered so that the IU could perceive this pointing action even if facing the other direction (see Figure 2). If this touch event occurred over an interactable object, this object would be *highlighted* with an orange contour. Finally, some puzzle actions could only be performed by the EU, such as dragging or re-sizing an object so it became graspable by the IU. Objects that were only interactable by the EU were shown in red (see Figure 6).

5.3 Puzzle Rooms

Six puzzle rooms were designed to explore the collaboration techniques above. These puzzles presented users with sequential challenges of increasing difficulty and no immediate instructions other than finding a way to open the room door:

Room #1: The simplest room, the VR user picks up a key but does not see the keyhole. The tablet user needs to rely on the VR user's vision cone (and associated hint mechanics) to identify an object it is able to rotate to reveal the keyhole. A door opens to the next room once the VR user places the key in the keyhole.

Room #2: Both users look for a red key that the tablet user needs to scale down (and thus turning it green) so it is graspable by its VR counterpart. The VR user then needs to place this key in a

keyhole after the tablet user has moved an object onto a pressure plate, opening the door to the next room.

Room #3: In this room the tablet user needs to flick a ball into a far away pressure plate. This plate is on an object the VR user can move around to facilitate the tablet user's task.

Room #4: This room combines the ideas above, requiring users to find a keyhole, enable a bridge for the VR user to cross, and requiring various objects to be moved onto different pressure plates.

Room #5: Similar to popular Portal games[1], in this room the VR user needs to place three objects in three pressure plates that are out of reach in elevated platforms. To reach them, the VR user needs to interact with a teleportation mechanism that brings her or him to another elevated platform that needs to be moved to the correct location prior to the teleportation.

Room #6: The final and arguably hardest room displays a series of columns that can be dragged around by the tablet user with the help of the hint technique. Upon staying in a pressure plate, the VR user enables a laser that needs to bounce from these columns in a such a way it reaches its destination and opens the final door. Participants completed the puzzle rooms in order, with the first two discarded as practice.

6 BI-DIRECTIONAL USER STUDY

In order to assess the perceived usefulness of the various pinpointing techniques implemented, particularly as the sole source of communication between IU and EU, we conducted a user study between pairs of participants attempting to complete the various puzzle rooms described above. We recruited 12 participants (5 female), all students at the local institution and aged between 18 and 25 years.

6.1 Experimental Design and Setup

Our study follows a within subject design with two conditions: VR and tablet; and each condition was explored in 36 trials (12 participants × three puzzle rooms). Both the VR and tablet applications were developed in Unity (version 2019.1): the VR experience was deployed on a first generation Oculus Quest via the Oculus SDK, and the tablet experience took place on a 11.6" tablet computer (1366×768) with the support of Lean Touch[2]. The communication between devices used the Open Sound Protocol (OSC) that resorted on extOSC[3], and we included various assets available online[4].

6.2 Procedure and Metrics

After completing the demographics, participants were assigned one of the two roles (VR or tablet) and given a brief introduction to the system and collaboration techniques. They were instructed not to speak to one another, and given no instructions other than trying to open the door they see in the virtual environment. The six puzzle rooms were completed in order, and after each one participants were asked to complete a brief survey where they selected the most useful technique, how hard was it to complete the task, and how much they felt the need to speak to do so (on a 5-point Likert scale where 5 corresponds to *"completely agree"*). Before starting a new puzzle participants would swap roles so that each would complete three tasks in each condition by the end of the study. After completing all puzzles participants completed the System Usability Scale (SUS) [2], the NASA-TLX [10], and provided some insights into their experience. This entire process took approximately one hour to complete for each pair of participants.

---

[1] https://www.thinkwithportals.com/ (Last accessed on 19-01-2021)

[2] http://carloswilkes.com/Get/LeanTouch

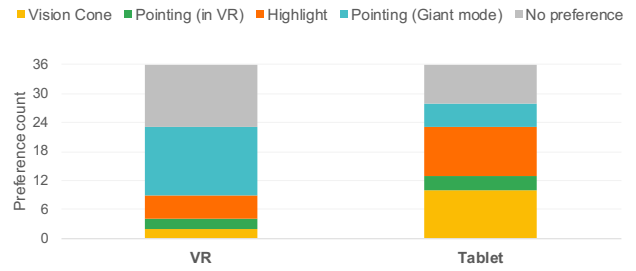[3] https://github.com/Iam1337/extOSC

[4] http://blitz3dfr.free.fr/

Figure 8: Self-report of which techniques were more useful in both VR and tablet settings (out of 36 trials each).

Table 1: Reported need to speak and task difficulty on a 5-point Likert scale across conditions (lower is better, std. dev. in brackets).

	VR	Tablet
Need to speak	3.39 (0.80)	3.11 (0.57)
Task difficulty	2.94 (0.69)	3.19 (0.56)

6.3 Results

All pairs of participants were able to successfully complete all puzzle tasks. Figure 8 highlights how participants rated each of the techniques in both conditions. The majority of participants in VR refered to the pointing technique triggered by their tablet counterparts (i.e. the giant fist) as the most useful technique (14), followed by the highlight technique (5). VR participants mentioned having no preference 13 times. On the other hand, tablet participants referred to the vision cone and highlight techniques as the most useful (10 each); followed by their own pointing technique (5). These participants mentioned having no preference 8 times. The least useful technique according to participants was the pointing technique in VR, mentioned only five times across VR and tablet participants.

Using a Wilcoxon Signed-Rank test, no significant differences were found for the perceived need to speak ($p = .212$) or task difficulty ($p = .091$) across conditions (see Table 1). These results highlight how consistent was the experience across conditions and their particular techniques. They also highlight how the tasks were perceived as somewhat challenging ($M = 3.07$), which could have led participants to report a moderate desire to speak to one another ($M = 3.25$). This result could have also been affected by our experimental setup: this was a co-located study where participants knew they could talk to one another if allowed. These challenges were furthered illustrated by a mean SUS score of 52.29 ($SD = 5.59$) and NASA-TLX of 11.70 ($SD = 0.58$).

Finally, participants described the experience as *"fun"* (P5, P6, P8) and *"enjoyable"* (P1, P2, P3, P7), and how *"the environment was simple (...) which allows the player to focus on (...) the communication with the other player, which was the best part of the experiment."* (P4). On the other hand, participants reported some frustrations over the communication (P3), particularly from the VR condition where a clearer pointing technique was suggested (P4).

7 DISCUSSION

In this work, we investigated how to support EUs in collaborating with IUs. Two uni-directional techniques were compared to a baseline alternative in the first study. The second study followed up with an evaluation of bi-directional techniques supporting bi-directional pinpointing tasks between IUs and EUs.

7.1 Performance

Techniques based on a disrupting element (such as a three-dimensional arrow or hand) resulted in faster completion times and were favoured by the majority of IU over less-disruptive techniques (e.g., highlight). P1 stated: *"Arrow is obviously the most efficient, you simply follow the arrow."*, while P6 said: *"[The] animation of the arrow made it easier to find the book."* Regarding the highlight P2 stated that *"Highlight needs [you] to explore, [but is] easy to understand; the Arrow needs low mind-load, is quick, and leads directly [to the object]; Voice requires high (mental) load."* The vision cone technique was also highly favoured by external tablet users in the second study, despite being more disruptive than the pointing animation on the IU avatar.

Based on these results, we think that both disruptive and highlighting techniques should become representative of standardised interaction paradigms for multi-user Cross-Reality applications. Their use should depend on the severity or urgency of the notifications. Disruptive techniques capture the attention of the user and are unlikely to be ignored, whereas highlighting techniques should be used for less urgent situations.

7.2 Cross-Reality Presence

While the analysis of Presence was not the explicit focus of our work, we asked participants in the first study to fill a custom questionnaire on their interaction with the EU. 62% of participants rated their interaction as "Excellent" (corresponding to a value of 6 or higher on a 7-point scale) which indicates that the interaction with them was considered to be overall positive. However, 25% of participants reported to not have experienced a high degree of immersion, with a score of 2 or lower. These results highlight how collaboration across realities can be beneficial, but might be perceived as being detrimental to one's feeling of immersion in their own "individual" reality. Future works should study the concept of Cross-Reality presence further.

7.3 Future Work

Pinpointing objects in another user's reality is one of the prerequisite steps before more complex forms of cross-reality collaboration can occur. In this work, we focused mainly on the modality through which users can focus the attention of users in another reality. Future work will need to investigate the effectiveness of the more disruptive techniques in contexts where users might be already engaged in other tasks or where multiple users are involved.

8 CONCLUSION

We have designed various novel Cross-Reality interaction techniques that support users at differing levels of immersion in pinpointing the attention of the other user on a specific object of the virtual environment. These techniques exemplify two varying levels of disruption. Results show that more disruptive techniques have a clear advantage in preference and in terms of time necessary for an immersed user to identify the object chosen by users who are not immersed. Conversely, when the pinpointing event is received by external users from immersed users, the former expressed a preference for a combination of disruptive and less disruptive techniques.

REFERENCES

[1] M. Billinghurst, H. Kato, and I. Poupyrev. The magicbook: a transitional ar interface. *Computers & Graphics*, 25(5):745–753, 2001.

[2] J. Brooke. Sus: a "quick and dirty' usability. *Usability evaluation in industry*, p. 189, 1996.

[3] L. Chan and K. Minamizawa. 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*, MobileHCI '17. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3098279.3098548

[4] E. F. Churchill and D. Snowdon. Collaborative virtual environments: An introductory review of issues and systems. *Virtual Reality*, 3:3–15, 1998.

[5] S. Ghosh, L. Winston, N. Panchal, P. Kimura-Thollander, J. Hotnog, D. Cheong, G. Reyes, and G. D. Abowd. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1447–1456, 4 2018. doi: 10.1109/TVCG.2018.2793698

[6] P. J. Gianaros, E. R. Muth, J. T. Mordkoff, M. E. Levine, and R. M. Stern. A questionnaire for the assessment of the multiple dimensions of motion sickness. *Aviation, space, and environmental medicine*, 72(2):115, 2001.

[7] J. Gugenheimer, E. Stemasov, J. Frommel, and E. Rukzio. 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*, CHI '17, pp. 4021–4033. ACM, New York, NY, USA, 2017. doi: 10.1145/3025453.3025683

[8] J. Gugenheimer, E. Stemasov, H. Sareen, and E. Rukzio. Facedisplay: Towards asymmetric multi-user interaction for nomadic virtual reality. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI '18, pp. 54:1–54:13. ACM, New York, NY, USA, 2018. doi: 10.1145/3173574.3173628

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

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

[11] J. Hartmann, C. Holz, E. Ofek, and A. D. Wilson. Realitycheck: Blending virtual environments with situated physical reality. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI '19. Association for Computing Machinery, New York, NY, USA, 2019. doi: 10.1145/3290605.3300577

[12] H. Ibayashi, Y. Sugiura, D. Sakamoto, N. Miyata, M. Tada, T. Okuma, T. Kurata, M. Mochimaru, and T. Igarashi. Dollhouse vr: a multi-view, multi-user collaborative design workspace with vr technology. In *SIGGRAPH Asia 2015 Emerging Technologies*, pp. 1–2. 2015.

[13] B. T. Kumaravel, C. Nguyen, S. Diverdi, and B. Hartmann. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. 2020. doi: 10.1145/3379337.3415827

[14] B. Laugwitz, T. Held, and M. Schrepp. Construction and evaluation of a user experience questionnaire. In *Proceedings of the 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering of the Austrian Computer Society on HCI and Usability for Education and Work*, USAB '08, pp. 63–76. Springer-Verlag, Berlin, Heidelberg, 2008. doi: 10.1007/978-3-540-89350-9_6

[15] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. In *IEICE Trans. Information Systems*, pp. 1321–1329, 1994.

[16] C. Nguyen, S. DiVerdi, A. Hertzmann, and F. Liu. Collavr: Collaborative in-headset review for vr video. In *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, UIST '17, p. 267–277. Association for Computing Machinery, New York, NY, USA, 2017. doi: 10.1145/3126594.3126659

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

[18] T. Piumsomboon, G. A. Lee, and M. Billinghurst. Snow dome: A multi-scale interaction in mixed reality remote collaboration. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, CHI EA '18, p. 1–4. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3170427.3186495

[19] D. N. Snowdon, A. J. Munro, and E. F. Churchill. *Collaborative Virtual Environments: Digital Places and Spaces for Interaction*. Springer-Verlag, Berlin, Heidelberg, 2001.

[20] A. Stafford, W. Piekarski, and B. H. Thomas. Implementation of god-like interaction techniques for supporting collaboration between

outdoor ar and indoor tabletop users. In *2006 IEEE/ACM International Symposium on Mixed and Augmented Reality*, pp. 165–172, 2006. doi: 10.1109/ISMAR.2006.297809

[21] C.-H. Wang, S. Yong, H.-Y. Chen, Y.-S. Ye, and L. Chan. HMD Light: Sharing In-VR Experience via Head-Mounted Projector for Asymmetric Interaction. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pp. 472–486. ACM, New York, NY, USA, 10 2020. doi: 10.1145/3379337.3415847

[22] H. Xia, S. Herscher, K. Perlin, and D. Wigdor. Spacetime: Enabling fluid individual and collaborative editing in virtual reality. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*, UIST '18, p. 853–866. Association for Computing Machinery, New York, NY, USA, 2018. doi: 10.1145/3242587.3242597