



Understanding Multi-user, Handheld Mixed Reality for Group-based MR Games

CARLOS AUGUSTO BAUTISTA ISAZA, Virginia Tech, USA

DANIEL ENRIQUEZ, Virginia Tech, USA

HAYOUN MOON, Virginia Tech, USA

MYOUNGHOON JEON, Virginia Tech, USA

SANG WON LEE, Virginia Tech, USA

Research has identified applications of handheld-based VR, which utilizes handheld displays or mobile devices, for developing systems that involve users in mixed reality (MR) without the need for head-worn displays (HWDs). Such systems can potentially accommodate large groups of users participating in MR. However, we lack an understanding of how group sizes and interaction methods affect the user experience. In this paper, we aim to advance our understanding of handheld-based MR in the context of multiplayer, co-located games. We conducted a study ($N = 38$) to understand how user experiences vary by group size (2, 4, and 8) and interaction method (proximity-based or pointing-based). For our experiment, we implemented a multiuser experience for up to ten users. We found that proximity-based interaction that encouraged dynamic movement positively affected social presence and physical/temporal workload. In bigger group settings, participants felt less challenged and less positive. Individuals had varying preferences for group size and interaction type. The findings of the study will advance our understanding of the design space for handheld-based MR in terms of group sizes and interaction schemes. To make our contributions explicit, we conclude our paper with design implications that can inform user experience design in handheld-based mixed reality contexts.

CCS Concepts: • **Human-centered computing** → **Mixed / augmented reality**; **Virtual reality**; **Collaborative interaction**; **Touch screens**; **Empirical studies in HCI**; **Computer supported cooperative work**; **Mobile computing**; **Mobile phones**; **Mobile devices**; **Tablet computers**; **Human computer interaction (HCI)**; **User studies**; • **Social and professional topics** → *Informal education*;

Additional Key Words and Phrases: Virtual Reality, Handheld VR, Handheld MR, Mixed Reality, Collaborative Virtual Environments, Social VR

ACM Reference Format:

Carlos Augusto Bautista Isaza, Daniel Enriquez, Hayoun Moon, Myounghoon Jeon, and Sang Won Lee. 2024. Understanding Multi-user, Handheld Mixed Reality for Group-based MR Games. *Proc. ACM Hum.-Comput. Interact.* 8, CSCW1, Article 197 (April 2024), 28 pages. <https://doi.org/10.1145/3653688>

1 INTRODUCTION

Virtual reality (VR) experiences are becoming popular among consumers due to technological advancements and the increasing affordability of VR head-worn devices (HWDs). While VR HWDs can offer high levels of immersion to users when experiencing virtual environments (VEs), VR

Authors' addresses: Carlos Augusto Bautista Isaza, Virginia Tech, Blacksburg, VA, USA, carlosaugusto@vt.edu; Daniel Enriquez, Virginia Tech, Blacksburg, VA, USA, denriquez@vt.edu; Hayoun Moon, Virginia Tech, Blacksburg, VA, USA, moonhy@vt.edu; Myounghoon Jeon, Virginia Tech, Blacksburg, VA, USA, myounghoonjeon@vt.edu; Sang Won Lee, Virginia Tech, Blacksburg, VA, USA, sangwonlee@vt.edu.



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2024 Copyright held by the owner/author(s).

ACM 2573-0142/2024/4-ART197

<https://doi.org/10.1145/3653688>

HWDs are poorly suited to some users. For example, children [9] and those who experience HWD-induced headaches or nausea [33] are discouraged from wearing HWDs. In addition, people who wear makeup [34], have coarsely textured hair [38], or wear prescription glasses [57] may not be willing to wear HWDs due to the discomfort or inconvenience that would result. Furthermore, certain activities that involve multiple dynamically moving users in a co-located space, such as recreational games, may not be compatible with the use of VR HWDs due to the risk of collision or potential injury [39, 52, 60].

In response to these challenges, handheld VR—using handheld devices as windows to view VR worlds—has been used as a solution for making VR content approachable to broader ranges of users [20, 24, 68]. This research thread effectively broadens the options available to accommodate users for whom VR HWDs are not an option, offering a new avenue to design for multiple users in VR and explore how expansive and dynamic this alternative VR experience can be. While previous works in handheld VR alluded to the potential of scalability in these systems [16, 20, 24, 25], larger-scale systems involving more than two users have yet to be studied.

While group size effects have been a recurring topic at CSCW [44, 58, 61, 65], we do not have a clear understanding of how group size in MR can affect the user experience. Understanding the effects of group size is significant because it changes the rewarding mechanisms and cognitive load that each individual can have [50, 69]. Group size is a particularly interesting topic in handheld VR, as users can maintain situational awareness in the real and virtual worlds simultaneously, turning any VR application into something approaching a mixed-reality (MR) experience [46]. This awareness, which is unavailable to HWD users, can invite more active, dynamic, and physical interaction, affording a sport-like game experience or socially constructed learning. For example, such systems will allow for novel types of group-based MR content, such as room-scale competitive games where tens of children run around holding tablets, or MR-based learning, where teachers teach a group of students with virtual content in social settings. Lastly, interaction design should also be considered in multiuser settings, as it can affect how groups of people interact with each other and use interactive systems [5, 19, 59].

This paper aims to investigate how group sizes and interaction methods can affect the overall user experience in a handheld VR game setting. We defined two research questions intended to serve as means of assessing the effects of group size and interaction methods on users' experiences in handheld, co-located MR experiences:

- (RQ1) How does the group size in a competitive game affect users' experiences in handheld MR?
- (RQ2) How do different types of interaction methods (proximity-based vs. pointing-based) affect users' experiences in handheld MR?

To answer these questions, we implemented MOMIS, a handheld MR environment and competitive game for up to 10 players. The object of the game is simple: claim more balloons in the virtual environment than the opposing team. The context of a group-based recreational game allows us to explore various types of interaction, incorporating both cooperative and competitive aspects of interactive systems. In addition, unlike alternative activities, such as team-based tasks, learning environments, or interactive demos, playing a game gives participants an immediate and explicit goal—to win the game—effectively motivating them to explore the physical space. We used the system to conduct a user study in which we varied group sizes within the game (1:1, 2:2, and 4:4) and the degree to which participants actively move, encouraged by two distinct interaction types: *Poke* (proximity-based interaction) and *Shoot* (pointing interaction at a distance).

Our results indicate that interaction types can affect social presence and perceived workload. In addition, we found that participants had a level of situational awareness in the physical world

sufficient to avoid physical contact with other participants even in the largest and most dynamic setting (4:4, Poke). However, individuals diverged in their preferences for game settings. Some participants prioritized group size preferences (smaller or larger), while others gave more weight to interaction type preferences (Poke or Shoot). Overall, participants felt positive about the dynamics that the game enabled in VR.

The findings of the study will advance our understanding of the design space of handheld MR in terms of group sizes and interaction schemes. To make our contributions explicit, we conclude our paper with design implications that can inform user experience design in handheld MR.

2 RELATED WORK

We identified three distinct areas that motivated our research: (1) the investigation of group size in co-located groupware; (2) co-located, handheld MR; and (3) the use of mobile devices to enable asymmetric, co-located VR systems.

2.1 The Effects of Group Size and Interaction Methods in Groupware

Existing literature on collaborative MR largely avoids delving into larger-scale collaboration involving more than two people. Researchers have given attention to the potential for large-scale collaboration, especially in the context of handheld MR; for instance, Rogers et al. identified a knowledge gap concerning multiplayer VR games with more than two users [56]. Similarly, Dagan et al. suggested “enable any number of players” as part of their design recommendations for mobile AR [14]. However, existing works that involve handheld VR have typically validated their studies with just one or two mobile device users during their evaluations [10, 16, 20, 24]. Authors also explicitly mention scaling interactions to accommodate larger groups as a direction to take up in future work. In practice, however, researchers tend to neglect the possibility of iterating larger-scale versions of the same applications. Consequently, we still lack an understanding of the effects of group size in co-located MR settings, despite the fact that many collaborative, handheld MR works can accommodate more than two users [18].

Investigating the effects of group size has traditionally been a prevalent topic at CSCW, especially when groupware technologies support novel, co-located collaboration [44, 61, 65]. Such studies have been particularly plentiful for technologies that encourage spatial interaction, such as tabletop displays or shared large-scale displays, and many researchers have also investigated these technologies’ impacts on task performance. Ryall et al. found that depending on the group size, people exhibited varying work strategies to achieve the same collaborative goal [58]. For instance, group size significantly influenced strategies for organizing shared resources [58]. Tang et al. discovered that closely shared perspectives were more important than avoiding disembodied visualizations of users in three-way remote communication [65]. Oftentimes, researchers study interaction methods or technological configurations that may exert an additional influence on interaction alongside group size effects. Researchers found that group size interacted with other design factors in groupware: how information was presented [59], how displays were configured [19], and the sizes of target objects [43]. These results suggest that designers must consider adjustments to their interaction schemes as group sizes increase. In our study, we aim to test whether the interaction between group size and interaction method (proximity- or pointing-based) can influence the user experience.

Lastly, the kinds of movements a game affords users for interaction can greatly affect the user experience [45]. In the case of “Brick”, a collaborative, handheld AR game, several insights were gained for designing room-scale collaborative games [10]. The researchers suggested that the interaction method should align with users’ existing mental models for familiar devices. On the other hand, multiplayer interaction in AR offers a broad design space for physical proximity and

exertion. Designers can consciously decide whether to emphasize these values or not. Another study proposed “encouraging touch through proxemics” for playful, co-located interaction [14]. In this regard, we chose to study the most intuitive interaction for the touch screen on a mobile device (Shoot, pointing-based) and another interaction type that can promote physical proximity and exertion (Poke, proximity-based). Our work further explores the design possibilities enabled by users’ even more rapid and active room-scale movement by incorporating competitiveness into our game design.

2.2 Handheld MR for Co-located Collaboration

Mobile device-only, co-located collaboration has been frequently explored in AR settings [17, 29, 30, 42, 53, 55]. Feng et al. conducted a survey of co-located AR collaboration between 2012 and 2022 and found that handheld AR is frequently used owing to its flexibility and in spite of its limitations (e.g., a narrow field of view and hands being occupied); half of the surveyed papers utilized hand-held displays [18]. However, using exclusively mobile devices for VR has not been frequently observed. The distinction between handheld AR and handheld VR should be based on whether virtual objects are positioned within the physical context or within a virtual environment (VE) [41]. Using a mobile device’s camera, on-screen visuals can represent the physical world in mobile AR with virtual objects overlaid, while the content on the screen can be entirely virtual in handheld VR¹. When users view VR content on mobile phones, the limited field of view available on small screens results in limited immersiveness, but increased awareness of the physical space and the co-presence of other users, given that users can see beyond the mobile device screens. In one study, the authors even referred to the mode of using monoscopic VR on mobile phones as “AR mode,” despite the on-screen visuals being entirely virtual [47]. The author’s perception of handheld VR might have been equivalent to AR due to the visual access to other actors and the physical environment available when using mobile devices to view VR content. Hence, the nature of handheld VR makes it challenging to distinguish between mobile VR and mobile AR. Müller et al. conducted a comparison of handheld VR and handheld AR in co-located and remote collaborative setups, finding no statistically significant differences between the two conditions in terms of task performance, preference, and workload [46]. The study did find individual differences; half of the participants demonstrated a preference for VR, while the remaining half expressed a preference for AR.

Compared to the long-standing and extensive history of literature in collaborative mobile AR [18, 54, 62], we have a limited understanding of the counterpart in handheld VR [46]. This gap suggests the existence of an underexplored space for larger-scale symmetric collaboration in VR using mobile devices. The ubiquity and adaptability of mobile phones and tablets facilitates access for many groups who are already familiar with the devices, as well as for those who cannot or would prefer not to wear an HWD; for example, individuals might be discouraged from wearing an HWD due to concerns about their health [9, 33] or personal appearance [34, 57]. While symmetric collaboration among HWD users outperformed asymmetric (HWD/mobile) and mobile-only conditions in object manipulation tasks, there was no significant difference in social presence [23]. This suggests that in contexts in which social engagement is central to the user experience (e.g., entertainment, learning, games), handheld collaborative VR may be an effective option. In our work, we explore the potential of handheld VR as an inclusive and social alternative to VR HWD-based collaboration.

¹In the context of this paper, we categorize early mobile phone-based stereoscopic HWD devices (e.g., Google Cardboard, which has been discontinued) [7] as a type of VR HWD. Mobile VR (or handheld VR) here refers to cases where a monoscopic view is displayed on handheld mobile devices. In literature, the term “head-held displays” (HHD) is often used to refer to mobile AR/VR [18, 22, 62].

2.3 Using Mobile Devices for Asymmetric Collaboration in VR

Many previous works have used mobile devices as an alternative to make VR accessible by allowing non-HWD users to view VR through these devices. The typical approach is to enable **asymmetrical** collaboration among HWD and non-HWD users, meaning that users use two distinct types of devices. ShareVR is one such example where a non-HWD user uses a mobile device that acts as a second viewport to the virtual environment (VE) [24]. The VE is also projected onto the floor from a top-down perspective, offering non-HWD users an extended physical representation of the VE. In VR Invite, Freiwald et al. developed a similar system to include older adults who may experience severe nausea when wearing VR HWDs [20]. TransceiVR took a different approach, allowing HWD users and non-HWD users to collaborate in co-located conditions. The system mirrored HWD users' perspectives and gave non-HWD users the ability to navigate their view temporally—that is, to rewind their perspectives [68]. WebTransceiVR extended TransceiVR by enabling online participants to see local users' views, essentially acting as cameramen [36]. Similarly, XRDirector used mobile devices to control camera movement in VR-based movie production, where VR HWD actors animate 3D characters [47]. Drey et al. conducted an evaluation of asymmetric collaboration in the context of learning, revealing that the use of mobile devices can lead to an increased cognitive load for teachers, yet also result in comparable learning outcomes [16]. FaceDisplay took a unique approach to sharing an HWD user's view with non-HWD users by mounting touch screens on the surface of a VR HWD. The displays on the surface of the VR HWD allowed people around the HWD user to see what the user saw and interact with the user [25]. Asymmetrical collaboration has also been applied to numerous commercial games, enhancing cooperative multiplayer experiences with unique asymmetric roles [1–4]. While these approaches vary, all of these works share the goal of making VR more inclusive of users who are unwilling or unable to wear HWDs. Our work aims to explore symmetric co-located collaboration using mobile devices exclusively, where users' roles shift from being secondary viewers to primary actors.

3 MOMIS: SYSTEM OVERVIEW

To explore handheld MR with support for multiple users, we created MOMIS, which stands for **M**obile-based **M**R for **I**nformal **S**ettings: a virtual environment (VE) platform that uses motion-tracked, six-degree-of-freedom (6DOF) mobile devices. The platform is designed to allow people to interact with a VE through a mobile device similarly to HWD users. A mobile user can look into the VE by using a mobile device that acts as a “window” into the virtual world, similarly to previous works [20, 24]. With this window, the mobile user can see the VE from the perspective that corresponds to their own position and orientation in the physical world. In this section, we present the technical and design details of the system. The platform was developed and improved over multiple iterations; the choices made may be useful for future researchers and practitioners creating handheld mixed reality systems. Figure 1 shows what the system looks like when multiple users are using the system at the same time.

3.1 Platform Specification

3.1.1 Hardware Setup. In this section, we introduce the hardware setup we used to enable 6DOF motion-tracked tablets. A VR-ready desktop computer runs SteamVR [64] and updates the positions of VR HWDs (if used) and tablet-mounted VIVE [13] trackers. Although we did not include VR HWD users in our study, asymmetric interaction is also feasible in MOMIS, as in previous works [20, 24]. The size of the virtual-physical environment in MOMIS is limited by the specifications of VIVE base stations, which support areas up to 10 meters by 10 meters [13]. We placed the VIVE base stations two meters above the ground to achieve better angles for tracking, using dedicated tripods.

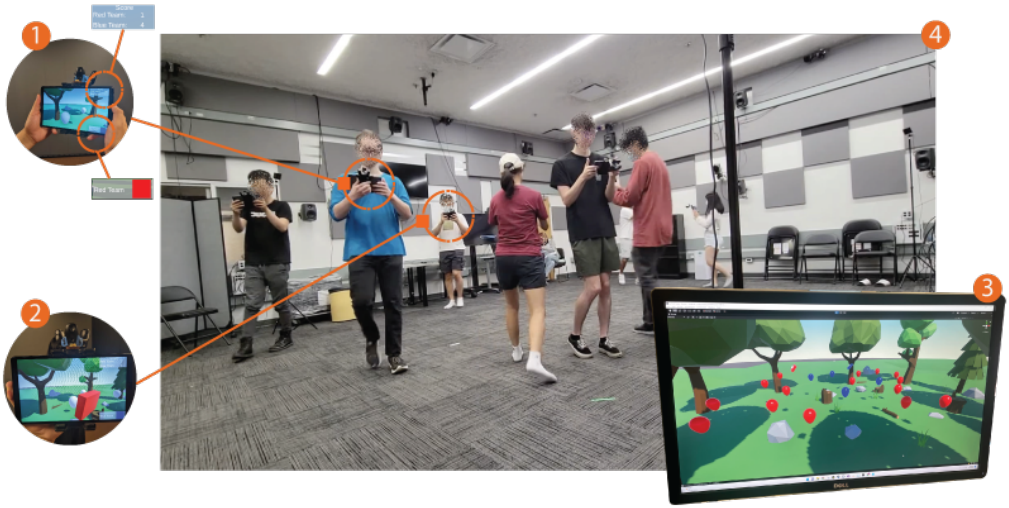


Fig. 1. A photo of a user study: (1) A user holding a 6DOF tracked device while viewing a VE. (2) A user viewing the VE and encountering another user's avatar. (3) An overview of an interactive game showing red and blue balloons; users interact with them to change the balloons' colors to their team's color. We tested two interaction types in the study: Poke (poking the balloons) and Shoot (shooting the balloons from a distance). (4) Eight users testing the prototype simultaneously.

To track the exact positions of the tablets, a VIVE tracker is mounted on top of each device. While we initially attached the VIVE trackers to the backs of tablets using Velcro tape, during the pilot study, we noticed that tablets could occlude the trackers' markers depending on how users oriented their tablets, especially when users oriented their tablets to interact with virtual objects at ground level. We changed the position to the top of the tablet using a commodity tablet mount to minimize occlusions by mobile devices or other users, especially given the amount of simultaneous users. We fabricated wooden attachments using laser cutters to mount each tracker to a tablet, as seen in Figure 2 (right). The VIVE base stations updated each VIVE tracker's position and orientation in real time.

The computer used to run the system was the Dell Alienware Aurora R12 [6] equipped with an 11th-generation Intel Core i7 11700F CPU running at 2.5 GHz and an NVIDIA GeForce RTX 3080 with 16 GB of memory. It concurrently ran the system at 90 frames per second and updated SteamVR at the same frequency. We chose Samsung Galaxy A7 Lite (8.7", 32 GB) tablets for their light weight and computing power, as after testing a few different tablet devices, we found that holding a tablet for an extended period of time can be physically exhausting.

3.1.2 Software Setup. The VIVE tracker mounted on a tablet dictates the position and orientation of the tablet user's viewing perspective in the VE. The position and orientation of the tracker relative to the center of the tablet was calculated and coded to correctly orient the tablet user's perspective. As users move while holding tablets, their positions in the VE are updated accordingly. To interact with others inside the VE, a user's position relative to others in the physical environment should match their virtual position. Therefore, to avoid desynchronizing users' virtual and physical locations relative to other users, locomotion methods that do not correspond to physical movements (e.g., teleportation) were not used.



Fig. 2. Tracker mounted on a mobile device: top view (left) and back view (right)

Because the VE necessarily corresponds to the physical environment where the system is installed, the virtual objects that users can interact with must be placed within the tracked and physically accessible area in the room. In MOMIS, we used a virtual space measuring 6 meters by 6 meters to match the scale of the development lab space and the scale of the user study space.

We developed a standalone mobile app in Unity [66]. The application's first screen presents users with a graphical interface with two consecutive screens: one where they can specify their tracker numbers, which are written on the physical devices (Figure 4-Right), and one to choose an interaction type (Figure 4-Left). After the brief configuration process, users can enter the VE and interact with the main experience.

To handle networking, MOMIS uses Photon Unity Networking (PUN) [51] to transmit information to the VR-ready computer via Wi-Fi to synchronize users and interactions inside the shared VE, and to update each tracker's position and orientation. The Photon server is instantiated upon creation by the host user (a research team member) on a VR-ready computer. The host user runs SteamVR, which concurrently updates and shares each tracker's position and orientation to all devices, allowing mobile device users to see the VE from where they stand. Interactions are shared through remote procedure calls using PUN. Users participating in the experience are visualized as tablet-shaped boxes so that users can see each others' locations through the screen as well.

3.2 MR Content: Competitive Multiplayer Game Experience

3.2.1 Game Design. We considered the following design goals in creating MR content for the MOMIS platform:

- Interaction must be intuitive; a novice user should be able to instantly learn how to use the system.
- The MR content must involve multiplayer components, such as collaboration or competition among users.
- The content needs to accommodate varying numbers of participants to allow us to investigate the effects of group size.
- The MR content needs to encourage users to dynamically and actively move within the virtual-physical space.

The principles specified above consider usability, practical needs in running a user study, the multiuser nature of the system, and the research questions we aim to answer. We sought to develop multiuser interactive content that is simple, yet engaging enough to create meaningful experiences through which to answer our questions. With these principles in mind, we decided to make a

simple, team-based game with differing interaction methods. It is worth noting that MOMIS is a *platform*, and the MR content presented in this section is just an instrument for our study. MOMIS could also be used for other types of content in other contexts.

In the game, players form two teams (red and blue) and compete against each other to accumulate points by interacting with virtual objects—more specifically, by changing the colors of virtual balloons. When the game begins, all the balloons are white. When a player interacts with a balloon, its color changes to match the player’s team color. Teams earn one point each time a balloon changes to the player’s team color. For example, if a player on the red team interacts with a white balloon, the white balloon will turn red, and the red team will score one point. Even after a team has claimed a balloon, the opposing team can still reclaim the balloon and change its color again. For example, the blue team can still earn a point by turning a red balloon blue. The game is thus continuous and competitive, as balloons can change colors repeatedly within the two-minute time limit. The experience requires users to move inside the virtual-physical world while interacting inside the VE. While the game is simple and straightforward, it can be influenced by how much players perform in consideration of others: other players’ locations, their anticipated behaviors, and collision avoidance for safety. This cognitive effort required of players is relevant to many theoretical constructs in collaborative system design (e.g., workspace awareness [26], spatiality [49], nonverbal communication [11], territoriality [61], and feedthrough [15]).

Other players inside the experience are represented by avatars with matching team colors. The avatars resemble floating flat boxes, patterned after the tablets and their orientations (note the red, tablet-shaped box in Figure 5-(Right)). A small cube on the back of the box represents the position of the tablet’s user. More realistic avatars with legs, arms, torsos, or heads would need to display additional information accurately to avoid interfering with users’ VE experiences. We did not attempt to visualize a full-body avatar because the assumed locations of players’ limbs inside the VE (in the absence of tracking data) may not be accurate; this may be an interesting topic for future research.

3.2.2 Interaction Methods: Poke vs. Shoot. One of the research questions that we set out to answer is how the extent to which users dynamically move affects the user experience: social presence, engagement, workload, and even safety issues (e.g., physical contact). To address these considerations, we developed two methods of interacting with virtual objects—Poke and Shoot—which differ in the required action and distance. For **Poke**, users directly make contact with a virtual object using their devices, as seen in Figure 3-Left. This method uses a simple collision detection algorithm. For **Shoot**, tapping the screen shoots a ball-shaped projectile or “bullet” capable of traveling up to two meters from the center of the tablet in the VE, as seen in Figure 3-Right.

The trajectory of the projectile in the Shoot interaction is a straight line orthogonal to the tablet plane at the time the user touches the screen. The projectile’s velocity was chosen to reflect a typical walking speed, such that players cannot save time by shooting balloons from a distance instead of walking to them. Initially, we tested ray casting as a third interaction method, analogous to the Poke method but using an infinitely long rod attached to the back of the player’s device in the virtual space. However, we realized that this made the game dynamics too chaotic; a player could potentially stand in one spot and continuously spin to find balloons, which would make the player dizzy. The speed and distance constraints in Shoot allowed users to interact with virtual objects at a short distance, but required them to adjust to the projectile’s trajectory. In both interaction types, a successful interaction (i.e., one that changes a balloon’s color) triggered a beep as an audio cue. In the case of Shoot, firing a projectile triggered a beep audio cue as well.

During the pilot study, we confirmed that the differences between Poke and Shoot affected how participants moved in the physical environment (see the supplementary videos). In the case of

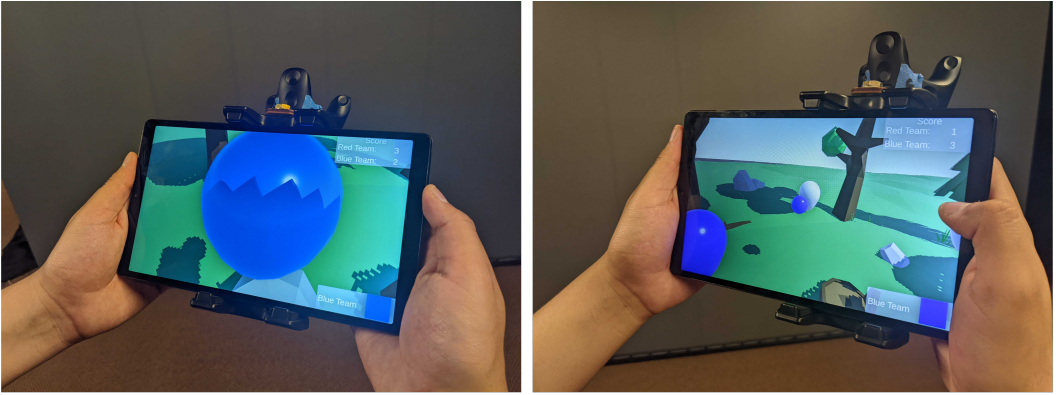


Fig. 3. Two interaction types developed for the game. (Left) A user “poking” a virtual object; (Right) A user “shooting” a virtual object.

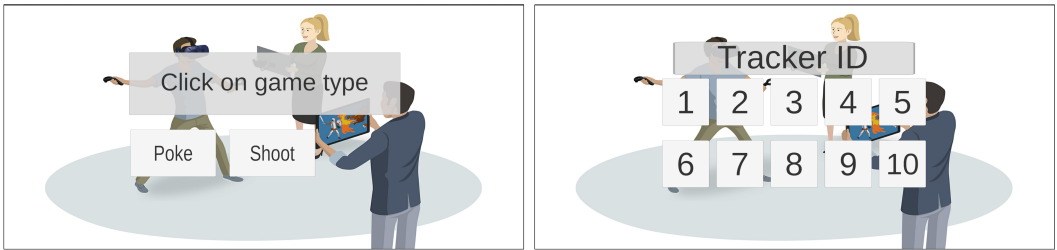


Fig. 4. Two lobby screens: (Left) The interface for choosing an interaction type; (Right) The interface for linking a tablet’s tracker to the system.

Poke, navigation and interaction require tablets—and by extension, their users—to move more or less constantly. Meanwhile, Shoot encourages users to locate target balloons, then stand still to aim, fire, and confirm the color change, resulting in less movement overall. These two interaction types allowed us to investigate the effects of people moving dynamically in handheld MR.

3.2.3 User Interface. The tablets run an Android application that renders a Unity scene. Apart from displaying the VE, the application also displays three UI modules: lobby configuration screens, a scoreboard, and a team color indicator. Before users join a game, they must choose either the Poke or Shoot interaction type (4-Left). After choosing an interaction type, users link their trackers to the system using the tracker’s ID number, as shown in Figure 4-Right.

Upon entering the experience, users see the VE and can explore it with their tablets, as seen in Figure 5-(Left). The scoreboard and team color indicators appear at the edge of the screen in translucent colors to minimize the occlusion of the VE. Thus, users can see which team they are on and what score each team has achieved.

4 USER STUDY

To explore the affordances and limitations of co-located, multiuser, handheld MR experiences, we designed the following user study. For reference, our research questions are presented once more below.

- RQ1: How does the group size in a competitive game affect users’ experiences in handheld MR?

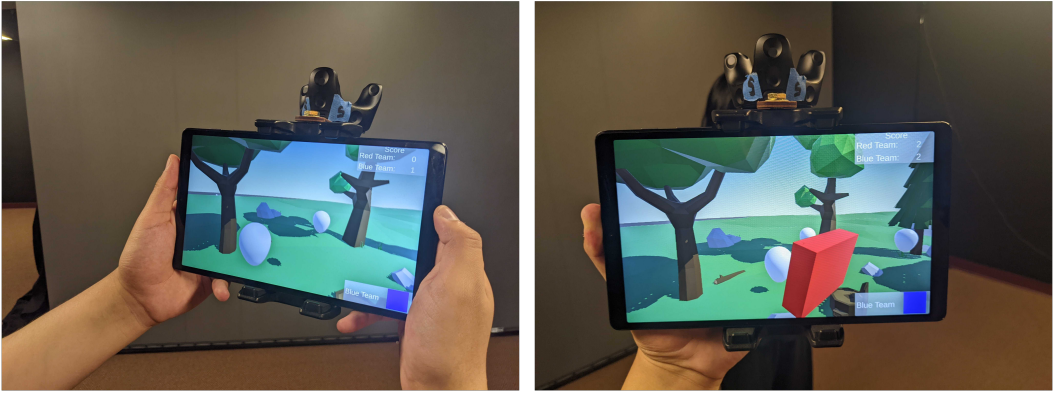


Fig. 5. Additional UI elements for the game. Left: The scoreboard overlay in the upper-right corner and the team color overlay in the lower-right corner help players understand the current state of the game. Right: Other players' tablets appear as pad-like boxes in their team colors. The red box is a visualization of another tablet that a user (behind the tablet) is holding.

- RQ2: How do different types of interaction methods (proximity-based vs. pointing-based) affect users' experiences in handheld MR?

4.1 Study Design

4.1.1 Independent Variables. To explore the effects of different experience conditions on the platform's affordances and user experience, we designed a 2×3 **within-subjects** study with two independent variables: **Interaction Type** with virtual objects and **Group Size**; that is, how many participants were playing the game at once. Each participant repeated the game six times, once in each of the six possible conditions.

We had two interaction type conditions; namely, *Poke* (proximity-based interaction) and *Shoot* (pointing-based interaction), as outlined in § 3.2.2; as well as three conditions for group size; namely, 1:1 (small), 2:2 (medium), and 4:4 (large). We scaled the group sizes by doubling, as opposed to linearly increasing the size, to explore a bigger scale without forcing participants to undergo an excessive number of trials. We did not include device type as a factor (VR HWDs vs. mobile devices) because VR HWDs are not viable options for certain populations, such as children (due to age restrictions) and individuals who experience motion sickness. Previous research has already compared VR headsets with mobile devices in a one-on-one setting, demonstrating a higher level of presence [24].

4.1.2 Dependent Variables. To measure participants' game experiences with different interaction types and group sizes, we wrote a post-game questionnaire asking about perceived presence, social presence, engagement, and workload for each experience condition.

- **[Presence]** We used a subset (three items) of Slater, Usoh, and Steed's presence questionnaire (SUS) [63] on (1) the sense of "being there" in the virtual environment, (2) the extent of perceiving the virtual environment as the dominant reality, and (3) the extent of perceiving the virtual environment as visiting a "place" rather than viewing images. We calculated a presence score by counting the number of 6 or 7 responses (on a seven-point scale) to the three questions, resulting in a number in the range [0, 3] as suggested by the original authors [63].

- **[Social Presence and Engagement]** We used the social presence module and the in-game module for social presence and engagement, respectively, from the Game Experience Questionnaire [31]. Among the three components within the social presence module, we used only six questions (all on a five-point scale) from the behavioral involvement component [24]. For engagement, we used two questions per component to measure each of the six components, yielding a total of twelve questions organized into six subscales: Competence, Flow, Positive Affect, Negative Affect, Tension, and Challenge. We discarded the Sensory and Imaginative Immersion components, as our goal was not to measure the quality of the game's plot or its visual appeal; indeed, the game had no storytelling component at all.
- **[Perceived Workload]** We used NASA-TLX [27] to measure perceived workload in terms of Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration, with values in the range [0, 100].
- **[Perceived Connectivity]** We investigated whether the increased group size and dynamic interaction could negatively impact **perceived connectivity** for various reasons; for instance, the tracking program on the desktop computer may slow down as the group size increases, or the tablet-mounted VIVE trackers may be occluded more in bigger groups or when players move more actively. We asked participants if they observed a loss of connection on their own devices and on others' devices on a scale from 0 (Never) to 4 (Always).
- **[Preference]** We asked participants to rank the six conditions in terms of preference, from Rank 1 (most enjoyable) to Rank 6 (least enjoyable).

We observed participants during the experiment to see if there were any **potential collision risks**. Participants also left open-ended feedback on what components made the experiences enjoyable or frustrating, and how these experiences compared with their previous VR experiences.

4.2 Procedure

Participants were recruited in groups of eight per session, as the large group size condition required four players on each of the two teams. Participants were briefed on the purpose and content of the experiment and confirmed their willingness to participate. Then, participants signed a consent form approved by the university's Institutional Review Board and were asked to complete a demographic questionnaire. Participants were assigned tracker IDs in the order of their arrival. If participants came in pairs, they were assigned to consecutive numbers to play one-on-one games with each other. Before beginning the experiment, participants were trained to use both interaction methods (Poke and Shoot) until they felt comfortable with them. Participants then experienced six different conditions with a mixture of two interaction methods and three group sizes: (a) 1:1 Poke, (b) 1:1 Shoot, (c) 2:2 Poke, (d) 2:2 Shoot, (e) 4:4 Poke, and (f) 4:4 Shoot. The order of conditions was counterbalanced for both independent variables, and each game experience lasted two minutes. After completing each game experience, participants completed a post-game questionnaire regarding their preferences and provided open-ended feedback.

4.3 Participants

We recruited 40 participants through online advertisements in university communities. Two participants, who arrived late, could not participate in all six sessions, so their data were excluded from the following analysis to maintain the within-subjects nature of the study. Those two participants were replaced by members of the research team in the sessions that took place before their arrival, so all sessions nonetheless had the correct participant counts. The remaining 38 participants (26 male, 12 female) had an average age of 25.79 (std. dev. = 5.01). Among the 38 participants, 27 had

prior experience with VR, and 26 had prior experience with multiplayer games. Each participant received a \$10 electronic gift card as compensation.

4.4 Data Analysis

We present the results of the statistical analysis of our quantitative and qualitative studies in this section. In the quantitative study, scores for Social Presence, Engagement, and Workload were analyzed using a 2×3 within-subjects experiment design (for interaction type and group size). We performed Shapiro-Wilk tests to determine if the data were normally distributed for all of the dependent variables. The data did not appear to be normally distributed, so we applied the nonparametric aligned rank transform (ART) process [71] and then ran a repeated-measures ANOVA for each dependent variable on the transformed values, using paired-samples t tests for post hoc analysis after a significant ANOVA. We used a Bonferroni correction to control for the multiple comparisons in this analysis. For example, a p -value of less than $\alpha = .05/3 = .017$ would be considered significant in a post hoc analysis of group size (1:1, 2:2, 4:4). In the specific case of presence, using the SUS presence questionnaire, we followed the procedure suggested by the authors to count the number of “6” and “7” responses to the three 7-point scale items. Therefore, the presence score was a value in the range $[0, 3]$ [63]. We used this score to perform a repeated-measures ANOVA on the ART-transformed values. We checked all the interactions between the two factors and there was no statistical significance between the two factors for all metrics, so we do not present the interaction term in the result tables. Instead, we present only the average and standard deviation per level of the two factors (Table 2) and the statistical test results (Table 1).

For the qualitative portion of the study, participants were asked to provide comments in the form of responses to open-ended questions. We used two questions for this purpose: (a) “What components made you enjoy the experience and what made you feel frustrated?” and (b) “If you had any previous VR experience, how would you compare this experience (e.g. VR headsets)?” We ran a thematic analysis and identified the most relevant topics and quotes to complement our quantitative results. Additionally, the researchers running the study conducted an informal observational study; we describe the most relevant observations in a separate section.

5 RESULTS

In this section, we outline our results in terms of each dependent variable: Presence, Social Presence, Workload, and Perceived Connectivity. We also present qualitative results from our own observations and the responses to open-ended items from the exit survey. The results of significance tests for each combination of a dependent and independent variable are summarized in Table 1. Averages and standard deviations are provided in Table 2, in the Appendix.

5.1 Presence

No factors had significant effects on Presence, although the figures for group size ($F(2, 74) = 2.90, p = 0.061$) and the interaction between group size and interaction type ($F(2, 74) = 2.57, p = 0.084$) suggest further research may be warranted. There was no significant effect found in terms of Interaction Type.

In our open-ended feedback, the majority of participants did not make any comments about immersiveness or presence in comparison to VR HWDs. Some participants (8/38) who had used VR HWDs admitted that VR HWDs are “more immersive” (P4, P10, P14) than mobile devices. A few participants (3/38) reported that they felt a high level of presence, stating that they were immersed in the game, or that the experience was comparable to (or better than) using VR HWDs.

- (P1) “*The games were very engaging. I almost forgot that I was in VR.*”

Table 1. F statistics and p -values from ART ANOVAs and significant pairwise comparisons. ($p < .05$ *, $p < .01$ **, $p < .001$ ***, Bonferroni correction applied for post hoc analysis on Group Size)

		Group Size		Interaction Type	
Presence		$F(2,74) = 2.90$, $p = 0.061$		$F(1,37) = 0.28$, $p = 0.599$	
Social Presence		$F(2,74) = 0.63$, $p = 0.533$		$F(1,37) = 17.86$, $p < 0.001$	Poke >Shoot ***
Engagement	Competence	$F(2,74) = 0.46$, $p = 0.636$		$F(1,37) = 0.85$, $p = 0.362$	
	Flow	$F(2,74) = 2.37$, $p = 0.1$		$F(1,37) = 0.05$, $p = 0.82$	
	Tension	$F(2,74) = 0.80$, $p = 0.455$		$F(1,37) = 11.44$, $p = 0.002$	Poke <Shoot **
	Challenge	$F(2,74) = 9.28$, $p < 0.001$	1:1 >4:4 ***	$F(1,37) = 0.06$, $p = 0.8$	
	Negative Affect	$F(2,74) = 9.28$, $p = 0.013$	2:2 >4:4 *	$F(1,37) = 0.97$, $p = 0.331$	
Workload	Positive Affect	$F(2,74) = 7.40$, $p = 0.001$	1:1 >2:2 *, 1:1 >4:4 **	$F(1,37) = 2.12$, $p = 0.153$	
	Mental Demand	$F(2,74) = 3.44$, $p = 0.037$	1:1 >4:4 *	$F(1,37) = 10.71$, $p = 0.002$	Poke <Shoot **
	Physical Demand	$F(2,74) = 2.51$, $p = 0.088$		$F(1,37) = 20.55$, $p < 0.001$	Poke >Shoot ***
	Temporal Demand	$F(2,74) = 1.59$, $p = 0.21$		$F(1,37) = 8.17$, $p = 0.007$	Poke >Shoot **
	Performance	$F(2,74) = 0.26$, $p = 0.772$		$F(1,37) = 0.01$, $p = 0.938$	
Perceived Connectivity	Effort	$F(2,74) = 5.15$, $p = 0.008$	1:1 >4:4 **	$F(1,37) = 1.45$, $p = 0.237$	
	Frustration	$F(2,74) = 1.05$, $p = 0.356$		$F(1,37) = 18.73$, $p < 0.001$	Poke <Shoot ***
Perceived Connectivity		$F(2,74) = 0.05$, $p = 0.955$		$F(1,37) = 0.24$, $p = 0.63$	

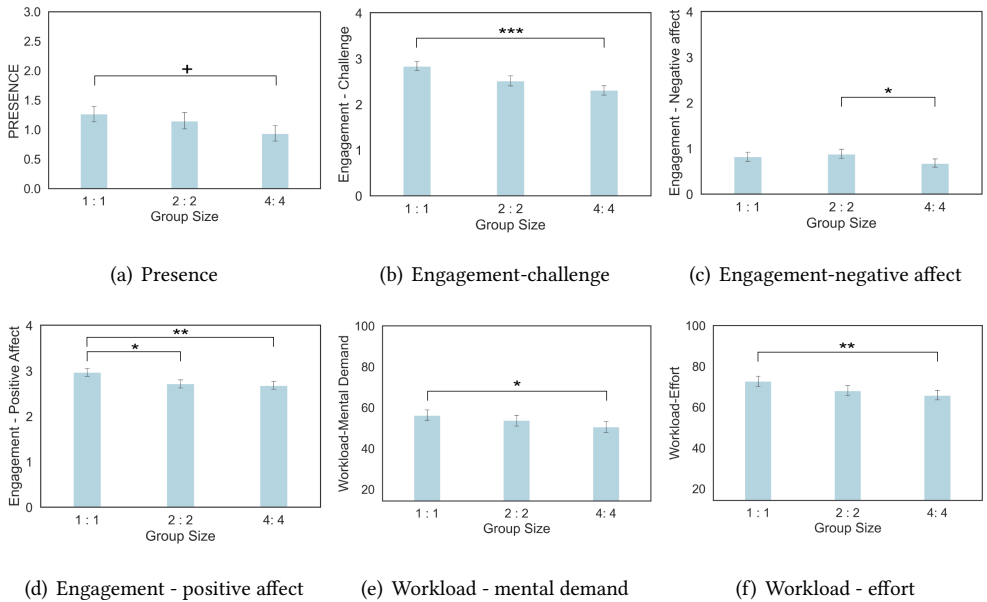


Fig. 6. Objective measures showing the main effects of the first independent variable, group size. Error bars represent standard error. Asterisks (*) indicate statistically significant differences between two conditions ($p < .05$ *, $p < .01$ **, $p < .001$ ***, with Bonferroni correction ($p < 0.05/3 = 0.0167$) applied for post hoc analysis on Group Size). A plus sign (+) indicates marginal significance at $p < 0.1/3 = 0.033$

- (P22) “I thought this was just as interactive as using a VR headset, which I was surprised about.”
- (P29) “Most fun I’ve had in a VR game.”

While it seems that only a few participants considered MOMIS impressive compared to alternative systems, the participants discovered other benefits in handheld MR, discussed later.

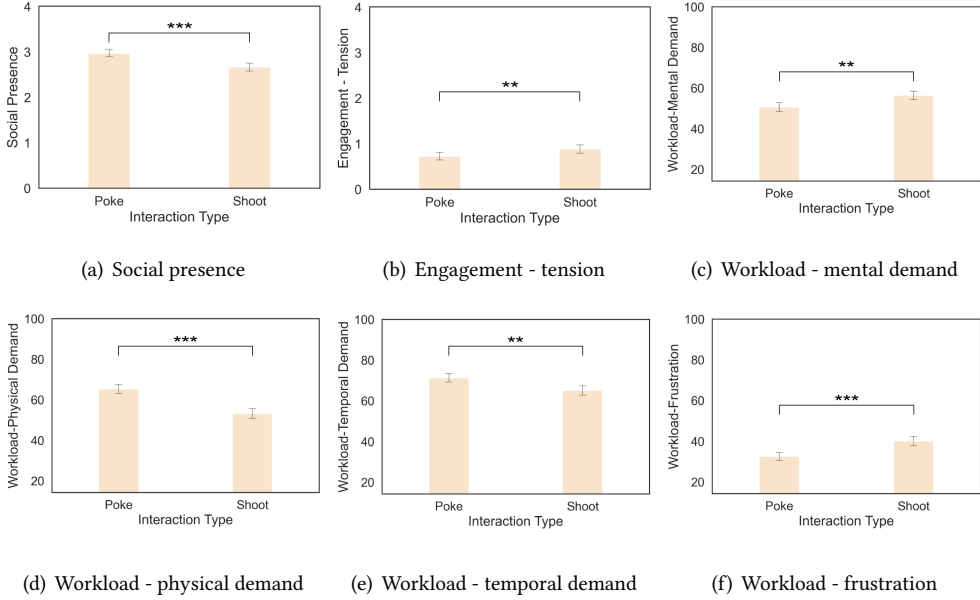


Fig. 7. Objective measures showing the main effects of the second independent variable, interaction type. Error bars represent standard error. Asterisks (*) indicate statistically significant differences between two conditions ($p < .05$ *, $p < .01$ **, $p < .001$ ***).

5.2 Social Presence

We found significant differences in social presence by interaction type. The result for Poke ($M = 2.97, SD = 0.82$) was significantly higher than that for Shoot ($M = 2.66, SD = 0.93, F(1, 37) = 17.86, p < .001$). The result indicates that using Poke may be more effective in fostering the sense of being with others. We could not find evidence that group size affected participants' perceptions of social presence.

Social presence in this context is not comparable with games or other VEs. In games or VEs, other users are visualized as avatars or game characters. In the case of MOMIS, users can see each other in the co-located space, and the VEs do not have full-body avatars as in other VR content. Therefore, even when they have to focus on on-screen content, the participants should have been able to maintain awareness of other participants through their peripheral vision, if not look directly at them.

5.3 Engagement

The effect of group size on challenge perception was statistically significant ($F(2, 74) = 2.65, p < .001$): *"I felt challenged, I had to put a lot of effort into it."* The post hoc analysis indicated that the 1:1 group size setup ($M = 2.84, SD = 0.90$) created the perception of a greater challenge than did the 4:4 setup ($M = 2.31, SD = 0.94, F(1, 74) = 18.46, p < .001$). No other pairs exhibited statistically significant differences. We anticipated that having more people could influence Engagement-Challenge in either direction; people may feel more challenged when there are too many people in the room, or less challenged when the game is collaborative. This result indicates that group size can modulate the extent to which users find gameplay challenging.

Group size also had effects on Engagement–Positive Affect and Engagement–Negative Affect. For the positive affect component of the GEQ in-game module (“*I felt content, I felt good*”), group size had a statistically significant effect ($F(2, 74) = 7.40, p < .01$). The post hoc analysis revealed that people felt more positive in the 1:1 setup ($M = 2.97, SD = 0.78$) than in the 2:2 setup ($M = 2.71, SD = 0.80, F(1, 74) = 8.80, p < .01$) and 4:4 setup ($M = 2.68, SD = 0.76, F(1, 74) = 13.01, p < .001$). For the negative affect component of the GEQ in-game module (“*I felt bored, I found it tiresome*”), group size again had a statistically significant effect ($F(2, 74) = 4.57, p < .05$). The post hoc analysis revealed that people felt more negative in the 2:2 setup ($M = 0.88, SD = 0.85$) than in the 4:4 setup ($M = 0.67, SD = 0.84, F(1, 74) = 8.59, p = .0045$).

Meanwhile, Interaction Type had a significant effect on Engagement–Tension. The Shoot interaction type ($M = 0.88, SD = 0.98$) produced more tension (“*I felt frustrated, I felt irritable*”) than the Poke interaction type ($M = 0.72, SD = 0.88$). However, the average score was relatively low—below 1—which indicates “slightly”. We did not identify any other significant effect on Engagement.

In the participants’ qualitative feedback, we found that most participants expressed how they were engaged with the game and the study for different reasons. In the open-ended feedback, we noticed that many participants (15/38) loved the experience of playing with other people, although some (13/38) did not enjoy the experience for the same reason.

- (P27) “*Multiplayer activity and actual physical movements were enjoyable.*”
- (P34) “*Way fun! Would love to play again! Enjoyed the ease of the way to gain points. [...] one of the reasons [my partner and I] don’t use our personal VR [head]set is because we only have one, and all the games are solo. Having people in the environment with me and completing a joint task is definitely my jam.*”
- (P22) “*I really enjoyed the 1 v. 1 because I only had to concentrate on one person as opposed to 2-3 other people.*”
- (P36) “*I enjoyed competing in 1v1 since the situation seemed [to be] in my control as opposed to basically random [in bigger groups].*”

We will discuss how people’s preferences widely vary depending on the group size in greater detail in § 5.6.

5.4 Workload

Workload, as measured by NASA-TLX, allows us to understand how two factors (interaction type and group size) affect the workload of their experience on six dimensions (Mental demand, physical demand, temporal demand, performance, effort, and frustration). Below, we report the dimensions for which we identified statistically significant effects.

5.4.1 Perceived effort is alleviated in a bigger group. We found significant differences in the effort scores of the different group sizes ($F(2, 74) = 5.15, p < .01$). The average score for 1:1 ($M = 72.75, SD = 21.4039$) was significantly higher than that for the 4:4 group size ($M = 65.79, SD = 20.48, F(1, 74) = 10.13, p < .01$). The result is consistent with the result for Engagement–Challenges found in the previous section. We did not identify any other significant differences in other dimensions by group size.

5.4.2 The design of the Shoot interaction may be mentally demanding and frustrating. Meanwhile, the interaction type produced different types of workloads. First, we found the Shoot interaction type ($M = 56.39, SD = 22.95$) to be significantly higher in terms of mental demand (Mental Demand: “How mentally demanding was the task?”) than the Poke interaction type ($M = 50.63, SD = 23.67, F(1, 37) = 10.71, p < .01$). In addition, we found that participants rated the frustration level (Workload–Frustration: “How insecure, discouraged, irritated, stressed, and

annoyed were you?”) of the Shoot interaction type ($M = 40.10, SD = 25.06$) as significantly higher than the Poke interaction type ($M = 32.58, SD = 20.69, F(1, 37) = 18.73, p < .001$). However, the overall frustration score was lower than approximately 40 in all conditions.

In the qualitative result, we found that this might have been relevant to the nature of the device being a viewing device and a pointing device simultaneously. While some people loved the Shoot interaction mode as is (P37: *“It is really intense. I liked the shoot part the most”*), some people (10/38) found the projectile’s trajectory difficult to understand.

- (P14) *“I felt the shooting mechanics were a bit off that made it a little frustrating.”*
- (P23) *“I enjoyed it but it was kind of tiring holding the tablet especially during the shoot games because I had to hold and aim.”*

Because participants had to aim correctly by holding the tablet at an angle orthogonal to the target, we determined that a pointing interaction ought to have visual aids, such as crosshairs or ray casting. Some other participants (5/38) pointed out how slow the projectile was. The slow projectile speed might have amplified frustration of Shoot interaction in view of the aiming interaction, which was already challenging on its own. Our design choices were based on the anticipation that near-instant speed and visual aids could decrease the amount of movement needed to interact with virtual objects. However, this type of design can increase the mental workload in such an experience.

5.4.3 **The Poke interaction may be physically and temporally demanding and frustrating.**

In the meantime, Poke had an impact on the perceived workloads of physical and temporal demand. On the NASA TLX, users rated the Poke interaction type ($M = 65.29, SD = 24.34$) as significantly more physically demanding (Physical Demand: “How physically demanding was the task?”) than the Shoot interaction type ($M = 53.26, SD = 24.95, F(1, 37) = 20.55, p < .001$). Similarly, in the temporal demand dimension of the NASA TLX (Temporal Demand: “How hurried or rushed was the pace of the task?”), we found the Poke interaction type ($M = 71.30, SD = 23.62$) rated significantly higher than the Shoot interaction type ($M = 65.16, SD = 24.27, F(1, 37) = 8.17, p < .01$).

It seems that people understood the game to be physically and temporally demanding by nature. No participants complained that playing in the Poke interaction type was specifically demanding. In fact, a few participants specifically mentioned a preference for the Poke interaction mode.

- (P11) *“Poke was fun, more than shooting. It was frustrating not being able to shoot fast as I wish.”*
- (P29) *“I found poke the most fun as I had to move my tablet around more in the VR space and as such found it more immersive.”*
- (P21) *“I really enjoyed just running around and playing competitive games.”*

P21’s comment offered a good summary of the affordances that MOMIS can provide for novel types of VR content. We also include a supplementary video in which readers can see how dynamically the participants walk and move during the game for Poke mode (see Supplementary Material).

5.5 Perceived Connectivity

We asked the participants about perceived connectivity to see if the group size and the interaction type had an impact on the performance of the system. Perceived connectivity did not differ significantly across conditions, neither by group size nor interaction type. On the item that states *“I observed a loss of connection on my device (e.g., screen freeze, delay): 0 (Never) to 4 (Always)”*, the median response was 1 ($M = 0.94, SD = 1.01$). This result indicates that a majority of users rarely perceived significant delays or disconnections, and we could not find any evidence that connectivity would be degraded by a particular interaction type or group size.

5.6 Preference

We asked all the participants to rank the six conditions in order from most enjoyable to least enjoyable. The two most preferred settings (those most frequently assigned Rank 1) were *1:1 Poke* and *4:4 Shoot*, which were ranked first by 11 out of 38 and 10 out of 38 participants, respectively. The distribution of ranks for each condition is shown in Figure 8 (left). The horizontal bar chart shows how many times each condition was ranked within the top 3 by shifting all the bars to match the borderline between Rank 3 and Rank 4 across the conditions. Figure 8 (right) further shows how each individual chose Rank 1 and Rank 6 conditions using a heatmap. For example, a total of 10 participants picked *4:4 Shoot* as their first choice, and 6 of those 10—the darkest cell in the matrix—picked *1:1 Poke* as their least preferred condition. We noticed a trend for those who prefer a smaller scale tending to pick the *4:4* setup as the least preferred and vice versa. We discuss this trend more in the Discussion section.

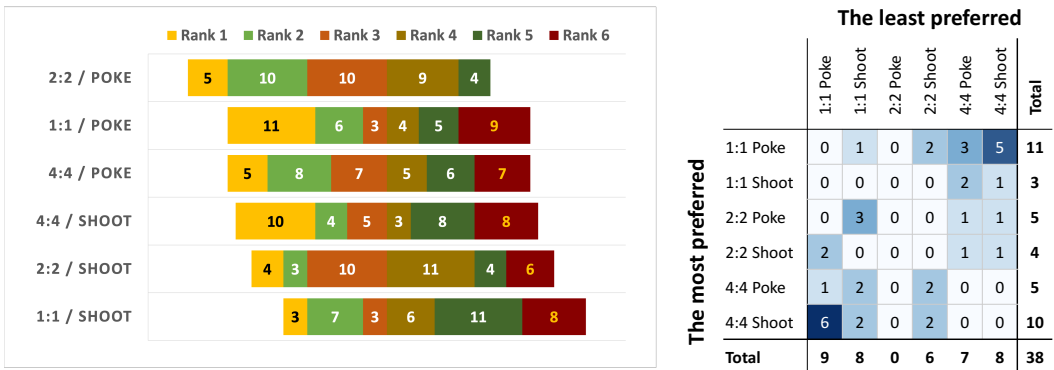


Fig. 8. Left: Rank distribution of each experiment condition. For example, there are five and zero participants who ranked 2:2 Poke as their first choice (Rank 1) and last choice (Rank 6), respectively. Right: Associations between preference rankings for different experiences. Rows represent the distribution of game conditions ranked as most preferred (Rank 1), while columns represent the distribution of game conditions ranked as least preferred (Rank 6).

5.7 Observation and Notable Findings

The authors continuously observed the participants' behavior during the study, as well as after the study (while reviewing the recording). We outline notable findings from our observations, as well as open-ended comments.

5.7.1 The participants played dynamically but safely. We observed that some participants moved dynamically during the game. We believe that this was possible because they had visual access to both the virtual world and the physical world. To gain an edge in competition, a few participants, especially while using the Poke interaction type, would run (see the supplementary video). However, while there were a few close calls, we saw no physical contact between participants. It seems that the participants understood the level of awareness needed to navigate the virtual world in relation to the physical world. One of the unique privileges of MOMIS is that users can have awareness in both the VE and the physical world, in contrast to HWD-based VR, where users are immersed in the VE to the exclusion of the physical world. P35's comment offers a good summary of this benefit.

- (P35) *“This is different because of the concrete spatial aspect, where you and others are in the same physical space. It required an increased awareness of actual surroundings while still immersed in the VR world.”*

Being able to move dynamically in the game was commonly pointed out as a unique component of MOMIS. The following are examples of responses to the item “How would you compare this experience with your previous VR experience of using VR headsets?” that exemplify how participants perceived MOMIS differently from VR HWDs.

- (P4) *“I think that headsets are more immersive simply because visually you can see nothing but the environment you are thrown into. However, for group activities I think this approach is better because with headsets it would be too clunky and I think participants would be more injury prone”*
- (P7) *“I would say that VR headsets would make me feel unsafe while playing the game since I [was] afraid of hitting the wall or someone else while can't seeing others around me. [...]”*
- (P27) *“I think VR headsets are more engaging. But for this type of activity, it can be risky because you cannot see other people.”*

Their awareness of both environments also affected their behaviors. Some virtual objects were close to the perimeter of the physical environment. When poking these objects, we observed participants holding their devices' views secondary to their attention to the physical world, using their physical spatial memory to avoid collisions with the walls. This required them to break their immersion in the application temporarily.

While they were able to move quickly, no one complained about nausea or motion sickness except one participant (P4, see below). This was another difference that some participants pointed out in the open-ended question.

- (P1) *“I usually don't like wearing a VR headset because of its weight. However, the tablet was so light that I almost forgot I was holding it. I also liked that I could use the wide space without motion sickness.”*
- (P18) *“Downside of VR is to make me feel dizzy but headset makes me dizzier but it doesn't.”*
- (P4) *“I felt motion sickness due to the lagging. Not tall-people-friendly enough.”*

Overall, the result suggests that MOMIS can be an inclusive VR option to accommodate people who experience nausea.

5.7.2 The limited field of view and weights. A few participants (3/38) reported that the limited field of view is a significant downside that might have impacted immersiveness for them. P36's comment below summarizes this point.

- (P36) *“Seemed like a smaller scale version of a VR headset experience, and was fun and immersive despite that, just not quite as immersive as real VR games”*

There were few opinions on the weight of the tablet. While one participant reported the tablets were “so light” (P1), other participants reported that the tablet was “a bit heavy to hold [with] single hand” (P9) and “[made] my arms feel tired sometimes.” (P7). However, given the short duration of the experience, we believe that we do not have evidence to conclude that a tablet can be considered a lightweight option.

6 DISCUSSION

Our work examines the impact of group size and interaction type on the design and implementation of multiuser, co-located MR experiences. We explored the system's design affordances using mobile devices with variance in terms of group sizes and interaction methods. Particularly, we documented our design choices through the iterative development process, including environment,

avatars, virtual objects, and audiovisual feedback. The user experience was evaluated with multiple relevant metrics useful for MR, groupware, and games: Presence, Social Presence, Engagement, and Workload. Our prototype revealed variance in some of these metrics depending on the group size and interaction type. What follows is a brief summary of our findings.

- Increasing group size alleviated perceived challenge, mental demand, and perceived effort needed, but reduced positive affect.
- Proximity-based interaction can facilitate social presence compared to pointing-based interaction.
- Proximity-based interaction was found to be more physically and temporally demanding than pointing-based interaction.
- Pointing-based interaction was found to be more mentally demanding and cause frustration.
- Participants were divided on their most preferred setups, with the two most popular setups being one that allowed for the most active movement (Poke, 1:1) and another that needed the least active movement (Shoot, 4:4).

We present and discuss the results organized by the system contribution and two research questions: Interaction Type (RQ1) and Group Size (RQ2).

6.1 (RQ1) Varying Group Size Can Modulate Challenge and Users' Affect.

The group size significantly affected a subscale of GEQ, Challenge (*"I felt challenged, I had to put a lot of effort into it"*). In particular, the participants felt that playing games in the 1:1 setting was more challenging than in a 4:4 setup. We had a similar result from the Workload-Efforts measurement as well; participants felt they had to put more effort into the 1:1 group setup compared to the 4:4 group. Being in a team makes it a group effort, which might have contributed to distributing efforts and challenges to the team and alleviating individual pressure. The diffusion of responsibility was shown to reduce cognitive effort and facilitate communication with those who shared responsibilities [50, 69]. In addition, given that the area of play and the number of balloons were controlled (held constant) across conditions, the perceived individual territory in larger-group settings may reduce the perceived cognitive load [61].

In the meantime, Positive Affect (GEQ) was significantly higher for 1:1 settings than 2:2 and 4:4 settings, suggesting a more positive gaming experience from playing 1:1 games compared to other setups. Challenges and Positive Affect may be correlated to some extent, especially in the context of game design, as pointed out by game researchers: what manifests through challenges embedded in playing games, such as failures, can encapsulate positive and negative emotion [32, 35]. This trend was also discovered in sports [69]; enjoyment decreased as team size increased. Another work showed that in a gamified learning setting, larger group sizes reduce students' engagement, as well as their effort and motivation, as individual contributions become less recognizable in larger groups [5]. Overall, the result suggests that increasing group size can diminish perceived challenges but also be effective in eliciting positive effects from gameplay with more agency. Group size had no significant effect on Presence, Social Presence, and other subscales of Workload and Engagement.

Accounting for individual differences, we observed a bimodal distribution of people in terms of their most preferred setup (Rank 1) regarding group size: a sizable group of users (15/38) ranked a 4:4 game setup (Poke: 5, Shoot: 10) as their most preferred setup, whereas a similar number (14/38) ranked a 1:1 game setup (Poke: 11, Shoot: 3) as their first choice. Naturally, fewer participants preferred the 2:2 setup the most (9/38). Interestingly, a majority of those who ranked 1:1 setups as their first choice ranked 4:4 setups as their least desired (11/14) (see Figure ??). Similarly, most of those who ranked 4:4 setups as their first choice ranked 1:1 setups as least desired (10/15). While we initially speculated about a "the-more-the-merrier" effect appearing in multiuser mixed reality, our study reveals that individuals may have different preferences regarding group size.

More research is needed to study the relationship between the number of players and gamers' experience and individual differences, as pointed out in previous works [28, 70]. One potential avenue for exploration in this regard is individuals' personality traits and skill levels (competence); research has found that physical capability and personality scores were able to predict the favorite game mode from among cooperative and competitive modes with 89.3% accuracy [48].

6.2 (RQ2) Interaction types affect social presence and types of workload in a game

The goal of RQ2 was to understand how the extent to which the participants dynamically move affects the user experience. We varied participants' movement by incorporating two interaction methods: Poke (proximity-based interaction) and Shoot (pointing-based interaction). Based on our observations, the two interaction types encouraged participants to play the game differently.

We discovered participants felt more social presence (GEQ) when they played with the Poke interaction type than with the Shoot interaction type. If our study had been conducted in HWD-based VR, we could have interpreted this sensation as participants feeling that they were together with others in the VE [12]. However, we believe this effect came from having visibility and audibility in a co-located space [11, 21]. We extend the literature that discusses the design space for physical proximity and contact [10] with the information that interaction design that encourages physical proximity and contact can promote social presence among other users in handheld MR.

In addition, we found how each interaction method can influence the different types of workloads. Our results indicate that participants felt more time pressure and felt physically tired using the Poke interaction. This result again is aligned with our observation that participants constantly navigated the space in the Poke condition. At the same time, it can serve as a manipulation check as far as the Poke interaction method's effectiveness in nudging the participants to move more dynamically relative to the other condition. Another component that may contribute to the physical demand is the need to constantly carry a mobile device. This result is also consistent with previous work; non-HWD users who preferred large displays (e.g., TVs) over mobile devices reported the convenience of virtual movement [20, 24]. Therefore, it is important to understand the extent to which designers would like to make a mobile VR experience physical and apply appropriate interaction methods depending on the target population, space, target experience duration, and context.

Lastly, using the Shoot interaction type increased perceived mental demand and tension/frustration. Aiming a target in the Shoot method certainly involves a perception and cognition task, and this may be more demanding than moving a tablet toward a target. The result that Shoot elicited less social presence than Poke while imposing greater mental demands is aligned with previous works that found that social presence was negatively correlated with mental demand [8, 67]. In addition, previous works noted that pointing-based interactions could be *inconspicuous*, less efficiently facilitating awareness of what a user's current action is [10]. Still, a considerable number of people (17/38) picked a condition with the Shoot interaction type as their most preferred setup.

Interestingly, the two most popular setups were two extreme options: 1:1/Poke ($n = 11$) and 4:4/Shoot ($n = 10$), the former being physical and strenuous, and the latter being cooperative and slow-paced. In light of the results we presented in the previous section, those who preferred 4:4/Shoot might have preferred the Shoot method because it is slow-paced and preferred the 4:4 setup because the setup alleviates perceived challenges and efforts. In contrast, those who ranked (Poke/1:1) first might enjoy the challenge and the effort demanded by the game under proximity-based interaction and the spatial freedom they have in a 1:1 setting. Previous works also showed a divide in users' preferences between cooperative and competitive modes for game-based rehabilitation, which could be accounted for by age and personality [48]. This finding again suggests the bimodal nature of people's preferences, which can be a challenge in designing mixed-reality content that is inclusive.

Therefore, understanding the nature of the target population and the user experience that the designers want in collaborative VR will influence the choice of interaction methods.

6.3 Applications of Larger-Group, handheld VR

Our study expands the literature of asymmetric collaborative VR, which mostly supports a dyad of an HWD user and a non-HWD user [24, 25, 68] by developing MOMIS, a handheld MR platform, and running an experimental study that involves larger groups of non-HWD users (up to 10 people in our case) in a safe and physically active MR experience. Not only does our study show that it is possible to develop engaging, fun, and physically active MR experiences using handheld technologies in co-located environments, but we also account for the effect of group sizes and interaction types on various aspects of the user experience. We believe that our study's findings will also apply to AR settings, as the difference between AR and VR is negligible in task performance [46]; rather, which setting to choose from among AR and VR seemingly depends on the nature of the virtual content and the physical environment (e.g., cluttered space vs. open space).

Our study reinforces the growing literature on making the MR experience social with co-located, multiuser VR experiences and can motivate new types of MR content that involve broader ranges of groups. Handheld MR can be useful for contexts in which users must be aware of both virtual and physical environments, including other actors and physical objects. In this work, we explored the multiplayer game context, which has already been frequently studied relative to multiuser, handheld MR [10, 14, 30, 37, 56]. Our findings can be applied to other contexts where gamification is applied, such as gamified learning environments, which typically include competitive components [40].

As we reviewed in 6.1 and 6.2, the findings from this study are consistent with previous works that did not involve games and competition. We argue that this is because there exists a common structure in handheld MR: multiple users in shared virtual objects. For example, this structure can create a competition-like relationship when multiple users are trying to interact with the same object. In this setting, proximity-based interaction will be more conspicuous than pointing-based interaction, promoting greater workspace awareness [26]. In addition, if the group size is larger in a collaborative setting, the diffusion of labor from a larger group size can alleviate users' perceived individual workloads, effectively lowering the mental barrier for novices.

7 DESIGN RECOMMENDATIONS FOR HANDHELD MR

The study results provide useful information to designers and developers (7.1, 7.2, 7.3). Furthermore, we document design insights that we gained from our iterative development process (7.4, 7.5). We present these as design recommendations below for those who aim to design multiuser, handheld MR experiences that avoid physical contact.

7.1 Leverage Users' Awareness of Both Virtual and Physical Worlds in Handheld MR

When designing an environment in a handheld MR setting, it is important to leverage users' awareness of both the virtual and physical worlds. Such awareness can effectively facilitate interactive user experiences that involve large groups of people moving dynamically. Users may not necessarily need full-body avatars of other users, as they can see each other beyond the mobile device screen.

7.2 Choose Interaction Methods Considering the Context and Goals

Based on our findings, UX designers can choose interaction methods in handheld VR settings to shape the user experience. On one end, they can create a competitive and physically demanding setting using proximity-based interaction methods, inviting intensive movement. Designers should also consider potential exertion associated with physically active interaction methods. On the other end, they can design less physically demanding pointing-based interactions. In addition, if

the goal of the UX includes facilitating social presence, employing proximity-based interaction is more desirable, as it encourages users to monitor both the VE and their physical surroundings. Similarly, pointing-based interaction can be subtle, reducing workspace awareness [10]. Therefore, UX designers can consider the context (space, nature of content, target duration, etc.) and the goal of an application in choosing which interaction methods to use.

7.3 Understand the Tension of Group-based Interaction

The effects of group size in the range under 10 (from 1:1 to 4:4) have varying impacts on engagement and workload in handheld MR. In general, running VR content with a larger group can alleviate perceived challenges and workloads, coming from competitive aspects of games or gamified UX, but there is more potential to elicit positive affect in a smaller group. In addition, participants' preferences can be divisive; there will be a group of people who prefer competitive, dynamic, and individual settings, while another group will prefer a cooperative, slow-paced environment [48]. One can take advantage of asymmetric roles, offering both dynamic and stationary roles in handheld VR to accommodate users with distinctive preferences [56].

7.4 Use Naturally Spatialized Auditory Feedback for Room-scale Interaction

While this was not part of the central research question, many users found it useful to have audio feedback when interacting with objects. Typically, a VR HWD sound system is for the user only, not others. In our design, we observed that providing sharp audio cues when interacting not only helped users identify whether they had succeeded in accomplishing a task, but also raised awareness of where other people were through sounds coming from their positions. Therefore, designers can include auditory displays to compensate for limited immersion and facilitate social presence and situational awareness.

7.5 Account for Performance Differences among Mobile Devices

Another practical insight we gained from the development of MOMIS is that the VR content developed for high-powered, VR-ready computers may not be rendered reliably on relatively low-spec tablets. During development, depending on the age and model of the mobile device in use, the mobile application sometimes ran slowly and would be considerably delayed in the environment and in frames per second. As a result, the research team made a mobile-only VE with fewer moving parts, fewer computationally intensive animations, a low-polygon environment, and few script updates. Thanks to this design, most people did not experience slowdowns on the tablet devices. Rendering content models with different fidelity and the same layout should be considered for translating VR content made for VR HWDs or supporting asymmetric interaction between VR HWDs and mobile devices.

8 LIMITATIONS

Our findings in this study may primarily apply to the specific context in which we developed our content: games that incorporate competitive and cooperative elements, or which otherwise have some characteristics in common (e.g., gamified, group-based learning). The proposed environment created a fast-paced and exciting user experience, allowing us to explore perhaps the most dynamic scenarios in which multiple users can consume an interactive virtual environment and interact with each other. However, different contexts with varying collaborative natures (e.g., work contexts) may have other factors influencing the results, including verbal communication and shared artifacts central to collaboration. Therefore, the study's findings will help us understand the design space of handheld mixed reality (MR), where people roam to experience VR content.

One component we did not consider was inclusivity as part of our experiment design and evaluation criteria (e.g., gender differences, including children, and diversifying group formation). While we believe that MOMIS, a multiuser, handheld MR, will be more effective than alternatives in including diverse types of users due to the device type used, we did not consider this factor in recruitment and metrics that we measured. For example, the studied participants were mostly young male adults (26 males out of 38, 18–39 age range with an average age of 25.8), meaning people outside this demographic group, such as children, teenagers, females, and middle-aged or older adults, were not properly represented. We believe that the best way to evaluate the inclusivity of the system is to conduct a field study, which we plan to carry out. We will discuss our plan to deploy the system in informal learning settings in §9.

This paper lacks an in-depth exploration of user behaviors and does not reveal the internal mechanisms and psychological processes of users' behavioral changes in response to the group size. This limitation is inherent in the study's quantitative design, which focused on evaluating the system using standard VR and gaming metrics. To address this weakness, we intend to conduct an observation-based field study with follow-up interviews as part of our future work.

9 CONCLUSION AND FUTURE WORK

Our work explores the experience of handheld mixed reality (MR) in multiuser competitive games. We developed MOMIS to conduct the experiment, which allows users to use mobile devices as windows into a virtual world to adapt to mixed reality, and which facilitates the use of social MR environments by including those who cannot wear head-mounted displays. This paper investigates the effect of group size (2, 4, and 8) and interaction methods (proximity-based and pointing-based) on the user experience. We found that proximity-based interactions positively impacted social presence and physical/temporal workload, while participants felt less challenged and motivated in larger group settings. Furthermore, individuals had varying preferences for group size and interaction type. This study informs user experience design in handheld mixed reality. It contributes to our understanding of the handheld MR design space regarding group size and interaction scenarios.

Our system offers the possibility of integrating non-HWD users into virtual environments (VEs) on a larger scale than just two users. It provides a platform for such users to actively participate, rather than being excluded. We plan to use MOMIS in informal learning settings to see if it can be applied to children's STEM learning materials. Informal learning settings like science museums are ideal places for us to see the inclusive and social benefits of MOMIS as children and their family members visit to enjoy collective learning experiences. We are currently developing STEM materials that will be displayed in MOMIS in collaboration with a local science museum (redacted for anonymity) to exhibit the work.

10 ACKNOWLEDGMENTS

This work is based upon work supported by the National Science Foundation under Grant No. NSF 2119011. The authors thank Dr. Phyllis Newbill, Dr. Todd Ogle, Dr. Doug A. Bowman, Dr. Benjamin Knapp, Dr. Lisa McNair, Ms. Mary Roberts-Baako, Dr. Chelsea Haines, and Shane Bennett for their feedback and contribution. Finally, we thank our user study participants for their time, effort, and feedback.

REFERENCES

- [1] 2016. Owlchemy Labs. Job Simulator (game).
- [2] 2016. VR The Diner Duo on Steam. https://store.steampowered.com/app/530120/VR_The_Diner_Duo/ publisher: whirly bird games.
- [3] 2018. Owlchemy. <https://owlchemylabs.com/blog/owlchemy-mobile-spectator-ar-spectator-camera>

- [4] 2021. IronWolf VR. <https://www.ionized.tech/> publisher: Ionized Studios.
- [5] Adnan Ahmad, Farrukh Zeeshan, Rutab Marriam, Alia Samreen, and Sohaib Ahmed. 2021. Does one size fit all? Investigating the effect of group size and gamification on learners' behaviors in higher education. *Journal of Computing in Higher Education* 33, 2 (Aug. 2021), 296–327. <https://doi.org/10.1007/s12528-020-09266-8>
- [6] Alienware. 2020. Alienware Aurora R12 Gaming Desktop | Dell USA. <https://www.dell.com/en-us/shop/desktop-computers/alienware-aurora-r12-gaming-desktop/spd/alienware-aurora-r12-desktop>
- [7] Ahmed Amer and Phillip Peralez. 2014. Affordable altered perspectives: Making augmented and virtual reality technology accessible. In *IEEE Global Humanitarian Technology Conference (GHTC 2014)*. 603–608. <https://doi.org/10.1109/GHTC.2014.6970345>
- [8] Huidong Bai, Prasanth Sasikumar, Jing Yang, and Mark Billinghurst. 2020. A User Study on Mixed Reality Remote Collaboration with Eye Gaze and Hand Gesture Sharing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3313831.3376550>
- [9] Jakki O Bailey and Jeremy N Bailenson. 2017. Immersive virtual reality and the developing child. In *Cognitive development in digital contexts*. Elsevier, 181–200.
- [10] Po Bhattacharyya, Radha Nath, Yein Jo, Ketki Jadhav, and Jessica Hammer. 2019. Brick: Toward A Model for Designing Synchronous Colocated Augmented Reality Games. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3290605.3300553>
- [11] Herbert H Clark and Susan E Brennan. 1991. Grounding in communication. (1991).
- [12] Carlos Coelho, Jennifer Tichon, Trevor J Hine, Guy Wallis, and Giuseppe Riva. 2006. Media presence and inner presence: the sense of presence in virtual reality technologies. *From communication to presence: Cognition, emotions and culture towards the ultimate communicative experience* 11 (2006), 25–45.
- [13] HTC Corporation. 2011. VIVETM | Discover Virtual Reality Beyond Imagination. <https://www.vive.com/>
- [14] Ella Dagan, Ana María Cárdenas Gasca, Ava Robinson, Anwar Noriega, Yu Jiang Tham, Rajan Vaish, and Andrés Monroy-Hernández. 2022. Project IRL: Playful Co-Located Interactions with Mobile Augmented Reality. *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 62 (apr 2022), 27 pages. <https://doi.org/10.1145/3512909>
- [15] Alan Dix. 1994. Computer supported cooperative work: a framework. In *Design issues in CSCW*. Springer, 9–26.
- [16] Tobias Drey, Patrick Albus, Simon der Kinderen, Maximilian Milo, Thilo Segschneider, Linda Chanzab, Michael Rietzler, Tina Seufert, and Enrico Rukzio. 2022. Towards Collaborative Learning in Virtual Reality: A Comparison of Co-Located Symmetric and Asymmetric Pair-Learning. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (New Orleans, LA, USA) (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 610, 19 pages. <https://doi.org/10.1145/3491102.3517641>
- [17] Maximilian Dürr, Marcel Borowski, Carla Gröschel, Ulrike Pfeil, Jens Müller, and Harald Reiterer. 2021. KiTT - The Kinaesthetics Transfer Teacher: Design and Evaluation of a Tablet-Based System to Promote the Learning of Ergonomic Patient Transfers. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (Yokohama, Japan) (CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 233, 16 pages. <https://doi.org/10.1145/3411764.3445496>
- [18] Shuo Feng, Weiping He, Xiaotian Zhang, Mark Billinghurst, and Shuxia Wang. 2023. A comprehensive survey on AR-enabled local collaboration. *Virtual Reality* (2023), 1–26.
- [19] Clifton Forlines, Chia Shen, Daniel Wigdor, and Ravin Balakrishnan. 2006. Exploring the Effects of Group Size and Display Configuration on Visual Search. In *Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work (Banff, Alberta, Canada) (CSCW '06)*. Association for Computing Machinery, New York, NY, USA, 11–20. <https://doi.org/10.1145/1180875.1180878>
- [20] Jann Philipp Freiwald, Sünje Gollek, and Frank Steinicke. 2021. VR Invite: A Project-Independent Smartphone App for VR Observation and Interactivity. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer International Publishing, Cham, 352–370.
- [21] Brian J Gajadhar, Yvonne AW de Kort, and Wijnand A IJsselstein. 2009. See no rival, hear no rival: The role of social cues in digital game settings. *Proceedings of CHI Nederland 2009* (2009), 25–31.
- [22] Eg Su Goh, Mohd Shahrizal Sunar, and Ajune Wanis Ismail. 2019. 3D object manipulation techniques in handheld mobile augmented reality interface: A review. *IEEE Access* 7 (2019), 40581–40601.
- [23] Jerônimo Gustavo Grandi, Henrique Galvan Debarba, and Anderson Maciel. 2019. Characterizing Asymmetric Collaborative Interactions in Virtual and Augmented Realities. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. 127–135. <https://doi.org/10.1109/VR.2019.8798080>
- [24] 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). Association for Computing Machinery, New York, NY, USA, 4021–4033. <https://doi.org/10.1145/3025453.3025683>
- [25] 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). Association for Computing Machinery, New York, NY, USA, 1–13. <https://doi.org/10.1145/3173574.3173628>
 - [26] Carl Gutwin and Saul Greenberg. 2002. A descriptive framework of workspace awareness for real-time groupware. *Computer Supported Cooperative Work (CSCW)* 11 (2002), 411–446.
 - [27] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in psychology*. Vol. 52. Elsevier, 139–183.
 - [28] Casper Hartevelde and Geertje Bekebrede. 2011. Learning in single-versus multiplayer games: The more the merrier? *Simulation & Gaming* 42, 1 (2011), 43–63.
 - [29] Anders Henrysson, Mark Billingham, and Mark Ollila. 2005. Face to Face Collaborative AR on Mobile Phones. In *Proceedings of the 4th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '05)*. IEEE Computer Society, USA, 80–89. <https://doi.org/10.1109/ISMAR.2005.32>
 - [30] Duy-Nguyen Ta Huynh, Karthik Raveendran, Yan Xu, Kimberly Spreen, and Blair MacIntyre. 2009. Art of Defense: A Collaborative Handheld Augmented Reality Board Game. In *Proceedings of the 2009 ACM SIGGRAPH Symposium on Video Games* (New Orleans, Louisiana) (Sandbox '09). Association for Computing Machinery, New York, NY, USA, 135–142. <https://doi.org/10.1145/1581073.1581095>
 - [31] W.A. IJsselstein, Y.A.W. de Kort, and K. Poels. 2013. *The Game Experience Questionnaire*. Technische Universiteit Eindhoven.
 - [32] Jesper Juul. 2013. *The art of failure: An essay on the pain of playing video games*. MIT press.
 - [33] Polyxeni Kaimara, Andreas Oikonomou, and Ioannis Deliyannis. 2021. Could virtual reality applications pose real risks to children and adolescents? A systematic review of ethical issues and concerns. *Virtual Reality* 26, 2 (2021), 1–39.
 - [34] Jolie Kerr. 2016. VR Headsets Are a Nightmare for the Makeup-Wearing Lady. <https://www.racked.com/2016/7/8/12097348/vr-headsets-makeup-removal>
 - [35] Pascaline Lorentz. 2014. Why failing in games is a positive aspect of play: A review of Jesper Juul's The Art of Failure. *Journal of Games Criticism* 1, 2 (2014), 1–4.
 - [36] Haohua Lyu, Cyrus Vachha, Qianyi Chen, Odysseus Pyrinis, Avery Liou, Balasaravanan Thoravi Kumaravel, and Bjoern Hartmann. 2022. WebTransceiVR: Asymmetrical Communication Between Multiple VR and Non-VR Users Online. In *Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) (CHI EA '22). Association for Computing Machinery, New York, NY, USA, Article 313, 7 pages. <https://doi.org/10.1145/3491101.3519816>
 - [37] Lissette López-Faican and Javier Jaen. 2020. EmoFindAR: Evaluation of a mobile multiplayer augmented reality game for primary school children. *Computers & Education* 149 (2020), 103814. <https://doi.org/10.1016/j.compedu.2020.103814>
 - [38] Arwa Michelle Mboya. 2020. The Oculus Go Wasn't Designed for Black Hair. <https://debugger.medium.com/the-oculus-go-a-hard-ware-problem-for-black-women-225d9b48d098>
 - [39] Mark McGill, Daniel Boland, Roderick Murray-Smith, and Stephen Brewster. 2015. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 2143–2152. <https://doi.org/10.1145/2702123.2702382>
 - [40] Nongluk Meekaew and Watcharee Ketpichainarong. 2018. An Augmented Reality to Support Mobile Game-Based Learning in Science Museum on Biodiversity. In *2018 7th International Congress on Advanced Applied Informatics (IIAI-AAI)*. 250–255. <https://doi.org/10.1109/IIAI-AAI.2018.00055>
 - [41] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, Vol. 2351. Spie, 282–292.
 - [42] D. Mogilev, K. Kiyokawa, M. Billingham, and J. Pair. 2002. AR Pad: An Interface for Face-to-Face AR Collaboration. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (Minneapolis, Minnesota, USA) (CHI EA '02). Association for Computing Machinery, New York, NY, USA, 654–655. <https://doi.org/10.1145/506443.506530>
 - [43] Neema Moraveji, Kori Inkpen, Ed Cutrell, and Ravin Balakrishnan. 2009. A Mischief of Mice: Examining Children's Performance in Single Display Groupware Systems with 1 to 32 Mice. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) (CHI '09). Association for Computing Machinery, New York, NY, USA, 2157–2166. <https://doi.org/10.1145/1518701.1519030>
 - [44] Meredith Ringel Morris, Kathy Ryall, Chia Shen, Clifton Forlines, and Frederic Vernier. 2004. Beyond "Social Protocols": Multi-User Coordination Policies for Co-Located Groupware. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) (CSCW '04). Association for Computing Machinery, New York, NY, USA, 262–265. <https://doi.org/10.1145/1031607.1031648>

- [45] Florian Mueller and Katherine Isbister. 2014. Movement-Based Game Guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) (CHI '14). Association for Computing Machinery, New York, NY, USA, 2191–2200. <https://doi.org/10.1145/2556288.2557163>
- [46] Jens Müller, Johannes Zagermann, Jonathan Wieland, Ulrike Pfeil, and Harald Reiterer. 2019. A Qualitative Comparison Between Augmented and Virtual Reality Collaboration with Handheld Devices. In *Proceedings of Mensch Und Computer 2019* (Hamburg, Germany) (MuC '19). Association for Computing Machinery, New York, NY, USA, 399–410. <https://doi.org/10.1145/3340764.3340773>
- [47] Michael Nebeling, Katy Lewis, Yu-Cheng Chang, Lihan Zhu, Michelle Chung, Piaoyang Wang, and Janet Nebeling. 2020. XRDirector: A Role-Based Collaborative Immersive Authoring System. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) (CHI '20). Association for Computing Machinery, New York, NY, USA, 1–12. <https://doi.org/10.1145/3313831.3376637>
- [48] Domen Novak, Aniket Nagle, Urs Keller, and Robert Riener. 2014. Increasing motivation in robot-aided arm rehabilitation with competitive and cooperative gameplay. *Journal of neuroengineering and rehabilitation* 11, 1 (2014), 1–15.
- [49] Gary M Olson and Judith S Olson. 2000. Distance matters. *Human-computer interaction* 15, 2-3 (2000), 139–178.
- [50] Richard E Petty, Stephen G Harkins, Kipling D Williams, and Bibb Latane. 1977. The effects of group size on cognitive effort and evaluation. *Personality and Social Psychology Bulletin* 3, 4 (1977), 579–582.
- [51] Photon. 2011. Photon Unity 3D Networking Framework SDKs and Game Backend | Photon Engine. <https://www.photonengine.com/pun>
- [52] Iana Podkosova and Hannes Kaufmann. 2018. Mutual Collision Avoidance during Walking in Real and Collaborative Virtual Environments. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games* (Montreal, Quebec, Canada) (I3D '18). Association for Computing Machinery, New York, NY, USA, Article 9, 9 pages. <https://doi.org/10.1145/3190834.3190845>
- [53] Lev Poretzki, Joel Lanir, and Ofer Arazy. 2018. Normative Tensions in Shared Augmented Reality. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 142 (nov 2018), 22 pages. <https://doi.org/10.1145/3274411>
- [54] Iulian Radu, Tugce Joy, Yiran Bowman, Ian Bott, and Bertrand Schneider. 2021. A Survey of Needs and Features for Augmented Reality Collaborations in Collocated Spaces. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 169 (apr 2021), 21 pages. <https://doi.org/10.1145/3449243>
- [55] Jun Rekimoto. 1996. Transvision: A hand-held augmented reality system for collaborative design. In *Proceeding of Virtual Systems and Multimedia*, Vol. 96. 18–20.
- [56] Katja Rogers, Sukran Karaosmanoglu, Dennis Wolf, Frank Steinicke, and Lennart E Nacke. 2021. A best-fit framework and systematic review of asymmetric gameplay in multiplayer virtual reality games. *Frontiers in Virtual Reality* 2 (2021), 694660.
- [57] Brittany Roston. 2016. Virtual reality headsets and glasses: when things get tricky - SlashGear. <https://www.slashgear.com/virtual-reality-headsets-and-glasses-when-things-get-tricky-11422243>
- [58] Kathy Ryall, Clifton Forlines, Chia Shen, and Meredith Ringel Morris. 2004. Exploring the Effects of Group Size and Table Size on Interactions with Tabletop Shared-Display Groupware. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) (CSCW '04). Association for Computing Machinery, New York, NY, USA, 284–293. <https://doi.org/10.1145/1031607.1031654>
- [59] Taylor Sando, Melanie Tory, and Pourang Irani. 2011. Impact of group size on spatial structure understanding tasks. In *2011 IEEE Pacific Visualization Symposium*. 107–114. <https://doi.org/10.1109/PACIFICVIS.2011.5742379>
- [60] Anthony Scavarelli and Robert J. Teather. 2017. VR Collide! Comparing Collision-Avoidance Methods Between Co-located Virtual Reality Users. In *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, Denver Colorado USA, 2915–2921. <https://doi.org/10.1145/3027063.3053180>
- [61] Stacey D. Scott, M. Sheelagh T. Carpendale, and Kori Inkpen. 2004. Territoriality in Collaborative Tabletop Workspaces. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA) (CSCW '04). Association for Computing Machinery, New York, NY, USA, 294–303. <https://doi.org/10.1145/1031607.1031655>
- [62] Mickael Sereno, Xiyao Wang, Lonni Besançon, Michael J. McGuffin, and Tobias Isenberg. 2022. Collaborative Work in Augmented Reality: A Survey. *IEEE Transactions on Visualization and Computer Graphics* 28, 6 (2022), 2530–2549. <https://doi.org/10.1109/TVCG.2020.3032761>
- [63] Mel Slater, Martin Usoh, and Anthony Steed. 1994. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 3, 2 (1994), 130–144.
- [64] Steam. 2016. SteamVR on Steam. <https://store.steampowered.com/app/250820/SteamVR/>
- [65] Anthony Tang, Michel Pahud, Kori Inkpen, Hrvoje Benko, John C. Tang, and Bill Buxton. 2010. Three's Company: Understanding Communication Channels in Three-Way Distributed Collaboration. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work* (Savannah, Georgia, USA) (CSCW '10). Association for Computing Machinery, New York, NY, USA, 271–280. <https://doi.org/10.1145/1718918.1718969>

- [66] Unity Technologies. 2019. Unity - Unity. <https://unity.com/>
- [67] Theophilus Teo, Gun A. Lee, Mark Billinghurst, and Matt Adcock. 2019. Investigating the use of Different Visual Cues to Improve Social Presence within a 360 Mixed Reality Remote Collaboration*. In *The 17th International Conference on Virtual-Reality Continuum and its Applications in Industry (VRCAI '19)*. Association for Computing Machinery, New York, NY, USA, 1–9. <https://doi.org/10.1145/3359997.3365687>
- [68] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. 2020. TransceiVR: Bridging Asymmetrical Communication Between VR Users and External Collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology (Virtual Event, USA) (UIST '20)*. Association for Computing Machinery, New York, NY, USA, 182–195. <https://doi.org/10.1145/3379337.3415827>
- [69] W Neil Widmeyer, Lawrence R Brawley, and Albert V Carron. 1990. The effects of group size in sport. *Journal of Sport and Exercise Psychology* 12, 2 (1990), 177–190.
- [70] Katherine A Wilson, Wendy L Bedwell, Elizabeth H Lazzara, Eduardo Salas, C Shawn Burke, Jamie L Estock, Kara L Orvis, and Curtis Conkey. 2009. Relationships between game attributes and learning outcomes: Review and research proposals. *Simulation & gaming* 40, 2 (2009), 217–266.
- [71] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Vancouver, BC, Canada) (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. <https://doi.org/10.1145/1978942.1978963>

A APPENDIX

Table 2. Average and Standard Deviation of Dependent Variables.

		Group Size			Interaction Type	
		1:1	2:2	4:4	Poke	Shoot
Presence [0,3]		1.3 (1.2)	1.1 (1.2)	0.9 (1.1)	1.2 (1.2)	1 (1.1)
Social Presence [0,4]		2.8 (0.9)	2.9 (0.9)	2.7 (0.9)	3 (0.8)	2.7 (0.9)
Engagement [0, 4]	Competence	2.6 (1.2)	2.8 (0.9)	2.8 (0.8)	2.8 (0.9)	2.7 (1)
	Flow	2.4 (0.9)	2.3 (0.8)	2.2 (1)	2.3 (0.9)	2.3 (0.9)
	Tension	0.9 (1.1)	0.8 (0.8)	0.7 (0.9)	0.7 (0.9)	0.9 (1)
	Challenge	2.8 (0.9)	2.5 (1)	2.3 (0.9)	2.5 (1)	2.6 (0.9)
	Negative Affect	0.8 (0.9)	0.9 (0.9)	0.7 (0.8)	0.8 (0.9)	0.8 (0.8)
	Positive Affect	3 (0.8)	2.7 (0.8)	2.7 (0.8)	2.8 (0.8)	2.8 (0.7)
Workload [0, 100]	Mental Demand	56.2 (22.4)	53.8 (23.2)	50.6 (24.6)	50.6 (23.7)	56.4 (23)
	Physical Demand	62.2 (26.6)	59.8 (23.9)	55.8 (25.4)	65.3 (24.3)	53.3 (24.9)
	Temporal Demand	71.1 (23.4)	68.4 (23.8)	65.2 (24.9)	71.3 (23.6)	65.2 (24.3)
	Performance	52.8 (28.4)	50.4 (25.6)	50.8 (24.8)	51.6 (26.2)	51 (26.4)
	Effort	72.7 (21.4)	68 (22.2)	65.8 (65.8)	69.6 (21.7)	68.2 (21.3)
Frustration		38.5 (24.7)	35.3 (22.3)	35.2 (22.8)	32.6 (20.7)	40.1 (25.1)
Perceived Connectivity [0, 4]		0.9 (1)	1 (1)	0.9 (1)	0.9 (1)	1 (1)

Received January 2023; revised October 2023; accepted January 2024