Examining the Effects of Immersive and Non-Immersive Presenter Modalities on Engagement and Social Interaction in Co-located Augmented Presentations

Matt Gottsacker
JPMorganChase
New York, NY, USA
University of Central Florida
Orlando, FL, USA
matt.gottsacker@gmail.com

Mengyu Chen
JPMorganChase
New York, NY, USA
mengyu.chen@jpmchase.com

David Saffo
JPMorganChase
New York, NY, USA
david.saffo@jpmchase.com

Feiyu Lu
JPMorganChase
New York, NY, USA
feiyu.lu@jpmchase.com

Benjamin Lee
JPMorganChase
New York, NY, USA
benjamin.lee@jpmchase.com

Blair MacIntyre
JPMorganChase
New York, NY, USA
blair.macintyre@jpmchase.com
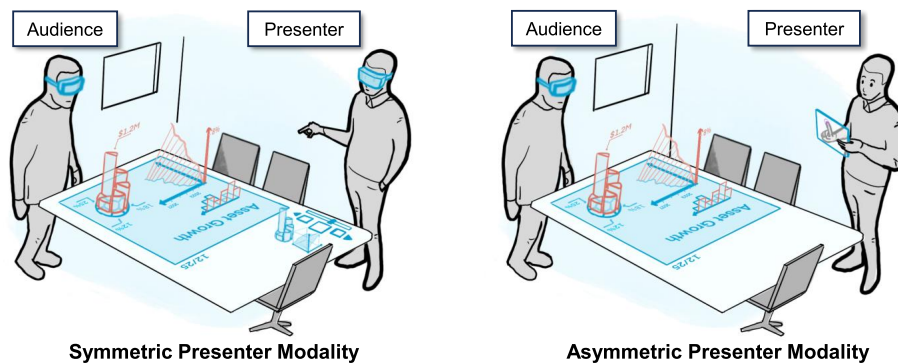
Figure 1: Illustrations of the immersive presentation system with symmetric and asymmetric presentation modalities. In the symmetric modality, the presenter uses a head-worn augmented reality (AR) display to control the presentation. In the asymmetric modality, the presenter controls the presentation with a tablet.

Abstract

Head-worn augmented reality (AR) allows audiences to be immersed and engaged in stories told by live presenters. While presenters may also be in AR to have the same level of immersion and awareness as their audience, this symmetric presentation style may diminish important social cues such as eye contact. In this work, we examine the effects this (a)symmetry has on engagement, group awareness, and social interaction in co-located one-on-one augmented presentations. We developed a presentation system incorporating 2D/3D content that audiences can view and interact with in AR, with presenters controlling and delivering the presentation in either a symmetric style in AR, or an asymmetric style with a handheld tablet. We conducted a within- and between-subjects evaluation with 12 participant pairs to examine the differences between these symmetric and asymmetric presentation modalities.

From our findings, we extracted four themes and derived strategies and guidelines for designers interested in augmented presentations.

CCS Concepts

• **Human-centered computing** → **Mixed / augmented reality**; **Collaborative interaction**; **User studies**.

Keywords

Augmented Reality, Augmented Presentation, Social Interaction, Engagement

1 Introduction

One of the greatest strengths of immersive technologies, such as augmented and virtual reality (AR/VR), is its ability to facilitate engaging and captivating experiences. Researchers, designers, and creative technologists have explored this strength for applications in education [42], healthcare [70], conferencing [57], entertainment [78], marketing [75], and more. Across these use cases, achieving higher levels of user engagement increases outcomes for learning, retention, productivity, relations, collaboration, and overall satisfaction. Among the engagement-focused immersive experience use cases, augmented and immersive presentations have emerged as topics of interest. Presentations — the delivery of information, concepts, or arguments — come in various forms and contexts, each tailored to specific purposes and audiences. Formal presentations use structured content and visual aids, while informal presentations are more casual and conversational, and interactive presentations involve audience participation. Regardless of the format, effective presentations require audience engagement, which can be enhanced through visual aids, storytelling, interactive elements, and real-time feedback mechanisms. The premise of *augmented presentation* is to leverage the affordances of AR/VR, such as spatial presentation, embodied interaction, and remote collaboration, to enable engaging, informative, and memorable presentations.

In co-located one-on-one presentation contexts (e.g., financial planning meetings, medical consultations, or creative design reviews) a single presenter controls the presentation flow for a single audience member who consumes and inquires about the presented information. In an augmented presentation setting, the presenter and audience may both use AR head-worn displays (HWDs) to have symmetric affordances for viewing and controlling the virtual presentation content. This symmetry could better enable presenters and audiences to have high group awareness (i.e., understanding what each other are attending to and referring to) to ensure effective communication, interaction, and understanding of the presentation topic and materials. However, even with AR HWDs, important social interaction cues such as eye contact could be diminished due to the wearer's face being obstructed [53], which can impair how people regulate conversation flow, express and understand emotions, and direct attention [1, 19]. Additionally, presenters may not want to wear an HWD for extended periods due to discomfort and fatigue [2], especially if they need to conduct multiple presentations back-to-back. In the early stages of this work, we observed this reluctance from stakeholders at our organization who regularly lead co-located one-on-one presentations. They emphasized the importance of social interaction in these presentations, which would likely be obstructed should they use an HWD. For these reasons, we explored asymmetric presenter modalities, such as a non-immersive tablet interface, to trade a presenter's level of immersion and group awareness in the symmetric presentation modality for possibly higher levels of social interaction. However, we chose to always have the audience view the augmented presentation with an AR HWD to prioritize their immersion in the experience.

In this paper, we present an exploratory user study with (n = 12) dyads examining how symmetric and asymmetric presenter modalities affect participants' presentation experiences regarding their interpersonal engagement and group awareness during the presentation. To conduct this study, we created a prototype presentation system with which an audience member with an AR HWD can view virtual presentation content, whilst a presenter can view and control the same presentation content using either a symmetric interface (i.e., also with an AR HWD) or an asymmetric interface (i.e., with a tablet). As asymmetry can result in reducing the common ground between users, our asymmetric interface provides additional features designed to maintain group awareness between the presenter and audience, including simulated mirrored audience views and the ability for the presenter to point to virtual content through those views. We conducted semi-structured interviews with participants about their engagement, group awareness, and overall experiences. Additionally, to study interpersonal engagement, we observed participants' social behaviors (e.g., how participants express and share social cues such as gestures and eye contact). For group awareness, we observed participants' referencing behaviors (e.g., how participants referred to presentation content spatially when directing each other's attention). Last, we collected questionnaire responses about presenters' cognitive load and presenters' and audience members' perceptions of engagement with the presentation (e.g., what was engaging) and their group awareness (e.g., how they referred to content or assessed their partner's attention) when using each interface. Our results suggest the benefits of immersive augmented presentation systems, strengths for each presenter modality, and open challenges for providing engaging presentation experiences.

The contributions of this work are as follows:

- An exploratory user study that examines the effects of symmetric and asymmetric presenter modalities on participants' social behaviors and group awareness.
- Four themes, derived from analyzing participants' user experience, around presentation techniques, spatial reference, social interaction, and interface impact.
- Design guidelines for researchers and practitioners interested in creating effective immersive presentation systems, with insights about trade-offs for selecting between symmetric and asymmetric presentation formats in co-located settings.
- The design of a novel prototype system for interactive, immersive presentations between physically co-located users including both 2D and 3D content and with symmetric and asymmetric presenter control modalities and interfaces.

2 Background

Our work is inspired by and grounded in three key areas of research: immersive storytelling and presentation, symmetric and asymmetric mixed-reality systems, and social interactions in cross-reality environments.

2.1 Immersive Storytelling and Presentation

Presenters often rely on captivating visual imagery to draw in audiences and increase their emotional and cognitive engagement with the narrative [39, 44, 93]. The desire to further bolster this engagement had given rise to a form of storytelling that closely blends the visual elements with the presenter's verbal statements and body movements, thus giving the illusion that the presenter is physically

"in control" and in sync with the displayed information. Ever since Hans Rosling's premiere example of this presentation style [68], research has strived towards making this form of storytelling more accessible and usable for laypeople. Said works have explored how interaction styles such as sketching [60], body gestures [30, 73], and speech [49] can further integrate the presenter's actions with visual elements in real-time. Yet, this form of storytelling treats the audience simply as a passive observer with little to no agency within the presented narrative. While acceptable for presentations to the masses, such as webinars, it is likely inadequate for more intimate co-located presentations where the audience's input and engagement are paramount (e.g., in sales pitches), especially when dialogue occurs between them and the presenter.

Prior research has demonstrated benefits of interactive elements in traditional 2D presentations in educational settings. For example, researchers have found that real-time audience polling systems for multiple-choice questions foster active participation and improve knowledge retention among medical professionals and patients [25, 84]. Additionally, systems that allow students to interact with presentation content more directly using their devices (e.g., by sketching or writing open-ended responses on their versions of the slides) enhance STEM students' learning and engagement [34, 69, 77]. Therefore, we sought to provide interactivity for the audience to help drive their engagement in the presentation.

Recent work has explored using interactive storytelling in AR contexts as well. Inspired in part by immersive theater performances involving spatial audience attention direction and multimodal interaction [95], Zhang et al. [99] introduced a system for creating and experiencing interactive immersive narratives using AR and IoT technology, which audience members found memorable and engaging. Additionally, Gottsacker et al. [21] proposed using head-worn AR to involve the audience in an augmented presentation directly. In it, the audience sees and can interact with visual elements in AR while the presenter narrates and controls the presentation via a tablet. While this approach encourages the audience to engage with the content for themselves directly, it still causes a visual separation for the audience between the input and output of the presenter's interactions, which the aforementioned presentation styles [30, 49, 60, 73] had sought to remove. They also did not evaluate the usability or effectiveness of the new presentation format.

Our work, therefore, seeks to compare the differences between having the presenter be physically and interactively "present" in AR with the visual content and audience in co-located augmented presentations (i.e., symmetric presenter modality) or be interactively separated from the audience through the use of a tablet (i.e., asymmetric presenter modality).

2.2 Symmetric vs Asymmetric Mixed Reality Systems

Mixed reality has been demonstrated for use in multi-user contexts beyond presentation, such as in education [6, 13, 56, 61] and remote assistance [27, 28]. Multi-user systems often have some form (or forms) of asymmetry in them as different users can have different experiences depending on the system's design [91]. There are multiple kinds of asymmetry involved in our scenario. First, there is asymmetry in presentation control because of the users' distinct

roles: the presenter has exclusive access to the control interfaces and controls which slides are shown and when. Additionally, because an AR HWD blocks part of the wearer's face, their facial expressions become difficult to observe by others, which may generally hamper social interaction in both of our presentation modalities. Our asymmetric condition also includes an input-output asymmetry: the presenter interacts with a 2D display to manipulate 3D content.

In mixed reality multi-user contexts, setups with "reality symmetry" are perhaps the standard approach, wherein all users utilize similar interfaces at the same point of the reality-virtuality continuum (RVC) [55, 80]. In a survey of co-located AR collaboration research, Radu et al. [64] emphasized how collaborators should be actively aware of each other's attention and emotions and be able to direct the other user's attention. Symmetric AR systems innately allow collaborators to have a shared awareness of what each other sees so that, in our case, the presenter can easily guide their audience throughout the presentation and the shared space [43, 50, 63, 64]. However, present-day HWDs block users' eyes, acting as a barrier to observing their mental states [23] and attention [52]. Perspective challenges also arise, making it difficult to reference areas of interest to one's peers—especially when the focus is on 3D content and visualizations [82]. To combat this, Chastine et al. [8] suggested providing multi-modal referencing techniques and shared viewpoints in collaborative AR environments.

On the other hand, asymmetric setups can have unique advantages in multi-user systems in general [91] and are gaining interest in XR contexts in particular as they provide "different means to visualize and interact with virtual content" [24], which can improve overall task effectiveness [24] while still facilitating a high degree of immersion to the non-HWD user [48]. Johnson et al. [38] also found that when suitable guidance cues are used, experts do not need to be in the immersive environment when referencing areas of interest—though this was mainly in a remote assistance context. For visualization, Reski et al. [66] found that having visualizations that share their position, orientation, and field of view of each user was able to support users' ability to make spatial references in an asymmetric setup. However, the incongruent interfaces and perspectives between users result in a diminished common ground [71], which can reduce workspace awareness [29] and thus, for example, make it more likely to interrupt a VR user disruptively [22, 41, 54]—or in our case, an audience member who is deeply engaged with the presentation content.

Quite relevant to our work, Drey et al. [13] investigated the effects of symmetric and asymmetric teacher interface modalities on communication and group awareness for co-located users in a pair-learning scenario. In the symmetric condition, both participants used VR HWDs and had avatars. In the asymmetric condition, the teacher used a tablet with teaching notes and a live spectator view of the student's VR view. We identify three key differences between our work and that of Drey et al.: (1) Drey et al.'s asymmetric design is suboptimal for social interaction, as the tablet user lacks an avatar in the virtual environment and has limited interactivity; (2) their setup represents a co-located VR environment where neither the other collaborator nor the physical surroundings are visible; and (3) their study focuses on paired learning, which differs from the presentation context. In contrast, our research explores symmetric and asymmetric AR setups for augmented presentations, allowing

both AR and tablet users to see each other and interact with the presentation content. In addition to group awareness, our study is particularly interested in the effects of different system configurations on users' social interactions in co-located AR scenarios.

Our work, therefore, considers all of these factors across both symmetric and asymmetric mixed reality systems and seeks to understand their influence on audience engagement, social interaction, and shared awareness in co-located augmented presentations.

2.3 Social Interaction in Cross-Reality

While recent events have led to a shift towards hybrid or fully remote meetings, people may still prefer in-person communication—especially to help establish more trust [89, 97] and better outcomes [4, 79] in high stakes discussions and negotiations. As such, an XR presentation system would likely seek to ensure that any conversation and social interaction between the presenter and their audience is as natural as possible, such as by leveraging behaviors standard in social settings as input modalities (e.g., speech, body gestures, eye gaze). For example, Dagan et al. [11] showed that designing co-located, symmetric mobile AR games to include shared experiences leveraging face-to-face social signals can foster social engagement. Moreover, presenters may eventually need to conduct regular immersive presentations in their day-to-day operations, which may prove fatiguing should an HWD be worn for long periods of time [2]. Interpersonal communication may also be negatively affected, such as by occluding the presenter's facial expressions and eye contact from the audience's view [53, 90]. For these reasons, it may be that asymmetric presentations, despite any possible downsides compared to symmetric ones, are ultimately more attractive and practical for the presenter than putting on a HWD. However, it remains an open question whether the presenter and audience occupying different ends of the RVC influence the overall social interaction and interpersonal engagement [59, 81].

Recent research highlights the complexities of fostering social engagement, cooperation, and connectedness when participants experience differing levels of immersion and interaction modalities. For games on traditional 2D screens, Harris and Hancock [32] have shown that designed interdependence in asymmetric cooperative games can improve players' perceptions of connectedness and social engagement over symmetric games. However, when comparing symmetric and asymmetric interfaces in a co-located 3D co-manipulation task using VR HWDs and AR tablets, Grandi et al. [24] found that asymmetric interfaces led to high cooperation but lower perceptions of mutual assistance. Additionally, asymmetric setups in which one user is immersed in XR can make XR users feel isolated and surrounding users feel excluded from the experience [26]. To address these challenges, researchers have provided additional channels for the non-immersed user to see and interact with the XR user's virtual environment to facilitate social interaction [26, 37]. Additionally, restoring social cues occluded by the XR HWD — such as displaying the XR user's eyes on the outside of the HWD — can improve social interactions [3, 7, 51, 52]. Additionally, Dey et al. [12] showed that sharing (typically invisible) physiological signals between VR users who cannot directly observe each other's social behaviors can enhance the quality of their social connection.

Our work examines how symmetric and asymmetric presentation styles influence interpersonal dynamics in co-located AR environments to help inform the design of augmented presentation systems that facilitate engagement and natural social interaction.

3 Prototype Design & Implementation

In this section, we describe the creation of our augmented presentation system prototype, which was used to investigate how symmetric and asymmetric presenter modalities affect the presentation experience of both the presenter and audience.

3.1 Augmented Presentation System Design Principles

We based the design of our augmented presentation system on related research on co-located AR and design goals Kumaravel et al. [87] derived from formative interviews about asymmetrical interactions between VR users and users of non-immersive displays. First, we aimed to support **independent exploration** for both users [87]. While it is a structured presentation where the presenter controls the main flow of the presentation, the audience can feel more engaged through interactivity [25, 34, 69, 77, 84]. To this end, we enabled both users to interact with the presentation content.

We also aimed to enable both users to make **efficient and direct spatial references** to presentation content [8, 64, 87]. This is required to enable the presenter and audience to discuss specific aspects of the presentation content. The symmetric setup naturally supports this goal as both users can use their hands to point to the shared AR content. The asymmetric condition, however, requires additional features to enable the presenter to make and understand references made in the AR space [38, 71] (e.g., mirrored views and attention guidance tools). Further, we aimed to provide **stable presentation content** for both users [87]. Simply, stability helps users make precise references to content when pointing to it. Again, this is straightforward for the symmetric modality because both presenter and audience view the AR content registered to the real world. To make precise references when using the tablet, the presenter needs a stabilized method of viewing the presentation content instead of through an unsupported hand-held view that is difficult to stabilize or through a view controlled only by the audience.

Last, we aimed to support **social interaction and co-presence** between users by providing social cues such as body language [64, 87]. Expressing and observing social cues is an important aspect of establishing and gauging interest during presentations [58]. The asymmetric and symmetric modalities have different strengths when it comes to meeting this principle, and investigating the trade-offs between them is indeed the objective of study in this paper. In particular, the symmetric modality supports a sense of shared context because both users interact in the same workspace. However, the AR HWD can obscure half of the presenter's face, blocking social expressions and potentially reducing interpersonal connection [53]. Therefore, the tablet may be better suited for supporting social interaction between users.
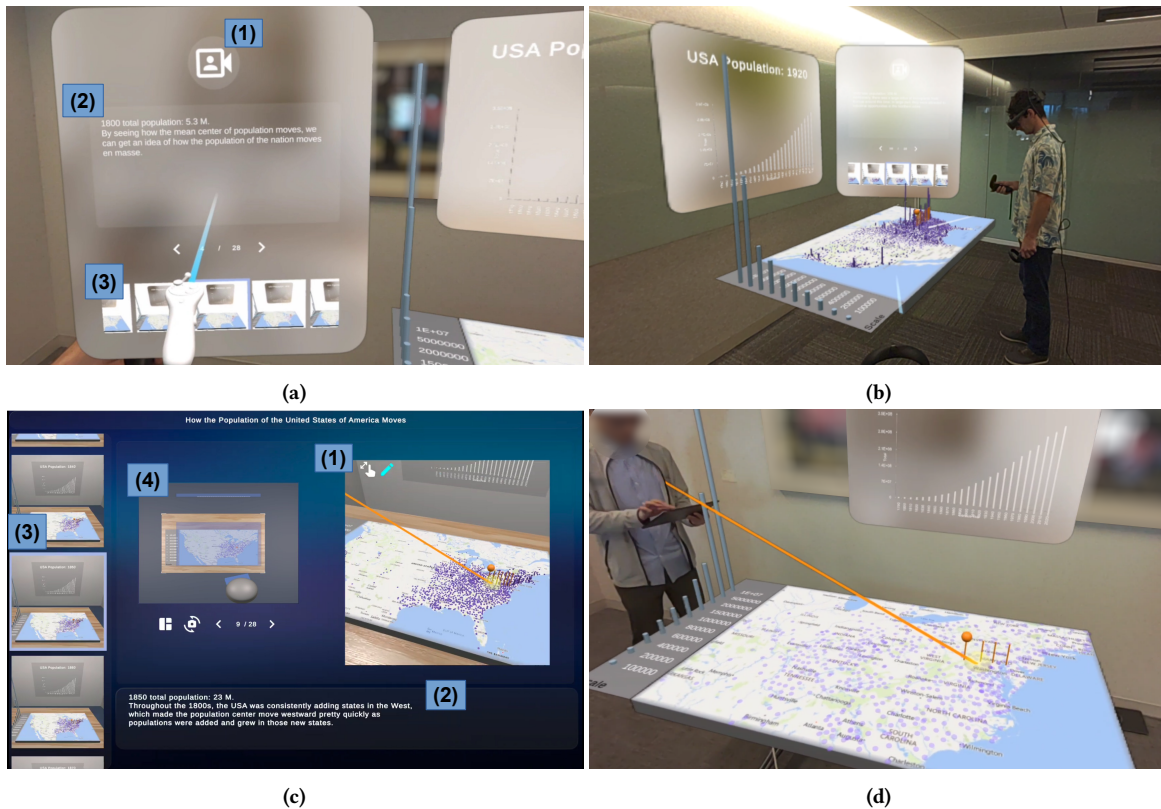
Figure 2: Screenshots from our augmented presentation system in use. (a) Symmetric presenter interface (close-up view). Features: (1) Button to toggle a mirrored view showing a simulated view from the audience's perspective. When activated, the simulated mirrored audience view appears directly above the pictured interface. (2) Speaker notes to guide the presenter on the presentation. (3) Scene switcher with a carousel view of previews of each scene in the presentation. (b) "Zoomed out" presenter's view of the presenter interface, the presentation content, and the audience. (c) Asymmetric presenter interface for the tablet. Features: (1) Simulated mirrored audience view. The presenter can tap on the virtual content to cast a virtual pointer ray. The presenter can press the hand icon button to switch to map interaction mode, allowing them to zoom the map by pinching with two fingers or pan the map by moving one finger. (2) Speaker notes. (3) Scene switcher. (4) Top-down view with simulated audience head to visualize the audience's position and orientation relative to the presentation content. The presenter can tap the presentation surfaces shown in this view (the map and the slide behind the map) to stabilize the simulated camera view so that it focuses on these objects. (d) Audience's perspective of the presenter using the virtual pointing ray when using the asymmetric presentation modality.

3.2 Presentation Interfaces

We designed and implemented separate interfaces for the audience and symmetric and asymmetric presenter modalities to best suit each device's visual and interaction affordances.

Audience Interface: The audience interface showed the presentation content situated in the real world through the AR HWD. The audience's controllers included standard ray interactors common in 3D environments [45]. Similar to other research involving user interactions with maps in immersive environments [98], the audience could pan the map by pointing at it with one of their controller rays, pressing the grip button, and moving the controller in the direction they wanted to pan the map. As an analog to zoom interaction on common map applications on mobile devices, they could zoom the map by pointing both controllers at the map, pressing both

controllers' grip buttons, and moving the controllers toward each other to zoom out or away from each other to zoom in.

Symmetric Immersive Interface: The symmetric presenter interface is shown in Figure 2a. It included the following features inspired by presenter interfaces for common desktop presentation software: a scene-switching scrolling view to allow the presenter to advance the presentation and preview scenes not currently shown, speaker notes to guide the presenter's speech about the current presentation scene, and a button to toggle a simulated mirrored view that shows the audience's current perspective of the presentation content. The presenter could point with the ray interactors on their controllers and interact with the map in the same ways as the audience. Based on findings from Piumsomboon et al. [62] showing

that persistent head gaze cues improved awareness in AR collaboration, the rays emanating from the audience's and presenter's controllers were always visible to each other to provide continuous interaction awareness cues.

Asymmetric Non-Immersive Interface: The asymmetric presenter interface is shown in Figure 2c. It included the same features as the symmetric interface, with additional features supporting group awareness between the presenter and the audience inspired by related research [64, 71]. To support a sense of shared space and provide awareness of the audience's location and attention, the tablet provided a simulated top-down view of the room, showing the presentation content and the audience's head position and orientation. The audience's simulated head in this view consisted of a simple capsule object with an attached cube to indicate the direction the audience is facing, similar to an avatar with a view frustum visualization (e.g., [62, 74]). Research on remote collaboration involving users who do not share the same task space has shown that sharing users' views can provide common ground and improve referencing communication [8, 17, 18, 38, 40, 71], and some systems have provided multiple views of the task space to provide more common ground [38, 65, 67]. Therefore, the tablet interface also provided a mirrored audience view panel, similar to the feature in the symmetric interface, which shows a simulated feed of the audience's view of the presentation content. Similar to effective attention guidance methods in remote-expert collaboration systems [47], the presenter could point to presentation content by tapping in the mirrored view panel. In the audience's 3D view, a virtual ray would connect with this pointing target (Figure 2d shows what this looked like from the audience's perspective). An alternative we had considered to support referencing for the presenter was to provide a tablet AR view of the scene that they could use to directly point to things with their real hands or with a virtual ray. In our initial design phase, we decided against this approach as we found it unwieldy for presenters to hold up the tablet to perform this interaction simultaneously with other presentation controls. Additionally, the presenter was able to select a presentation surface in the top-down view, and the mirrored view panel showed a camera that was stabilized and focused on that surface. This was designed to allow the presenter to point to specific parts of the presentation content or manipulate the map more precisely (i.e., without being affected by the audience's movements in their live feed).

To maximize interaction learnability and familiarity, the presenter could interact with the map with common map interactions on touchscreen devices: they could pan the map by dragging one finger on the map, and they could zoom the map in or out by using two fingers and separating or pinching their fingers, respectively.

3.2.1 Hardware and Setup. We built the presentation system using Unity version 2022.3.4. The system consisted of four applications: an audience application that ran on an AR HWD, a symmetric presenter application that ran on an AR HWD, an asymmetric presenter application that ran on a tablet, and a server application. The audience application and the symmetric presenter application were run on Dell Precision 7680 laptops each equipped with an Intel i9 CPU, an NVIDIA RTX 4000 Ada Generation Laptop GPU, and 64

GB of RAM. Each laptop was connected to an HTC VIVE XR Elite HWD with a VIVE Streaming Cable. The asymmetric presenter application was run on a Samsung Galaxy Tab S8+. The server application relayed networking messages between the presenter and audience to synchronize presentation scene data and user position data. Networking was implemented with Mirror Networking[1].

4 User Study

We believe augmented presentation has the potential to benefit the experience and outcomes of co-located one-on-one presentations. To further examine this premise, we constructed and conducted a user study using our system for augmented presentations. This work examines the effects of augmented presentation and symmetric or asymmetric presenter modalities on engagement, group awareness, and social interaction. This section details the study design, implementation, and execution of our within- and between-subject user study, designed to satisfy this goal.

We designed a within- and between-subject user study in which participant pairs utilized our augmented presentation system to experience or present a pre-authored presentation. We collected qualitative behavioral data, self-reported questionnaires, and semi-structured participant interviews. The presenter control interface was designed within subjects where each presenter presented twice, once using the symmetric immersive interface for one and the asymmetric non-immersive interface for the other. The order in which the interface was used first was counterbalanced among presenters. This design was chosen to allow us to glean insights from presenters on the differences and preferences between control interfaces. The audience experience was designed between subjects. Each audience participant saw the presentation once with a presenter using one of the two control interfaces. This design was chosen to allow us to observe differences in audience behavior and responses with each control interface.

4.1 Presentation Content

Using our system, we designed and implemented a narrative data-driven presentation incorporating static and interactive elements in 2D and 3D, as both are prevalent formats of information presentation and consumption in today's presentation/exhibition experiences [14, 20, 83, 88]. We chose US census data as the presentation topic, as it is well suited for spatial presentation, visualization, interaction, and is straightforward and relevant for many people even without prior background or expertise. Using this data, we created a presentation about the history of population growth and migration in the US from 1790 to 2020, inspired by a narrative visualization video by Danielle Kurtzleben and Vox[2]. The presentation was divided into slides for each decade of our time range. Each slide in the presentation contained a map visualizing this data in 3D, and a 2D bar chart visualization of the total population by decade. Each slide also had speaker notes as a reminder to the presenter of the changes in the data across each decade. The map was spatially registered to a real table. The population data was visualized on this map with 3D cylinders for each county, population values

---

[1]https://github.com/MirrorNetworking/Mirror
[2]https://www.youtube.com/watch?v=OTaYKEumwc8

being encoded by height. Furthermore, orange cylinders encoded the population mean center[3] for each decade to further emphasize the shift in the US population over time.

4.2 Measures

We used a combination of self-report questionnaires, video recordings of experiment sessions, and semi-structured interviews to examine participants' presentation experiences in terms of engagement, group awareness, social interaction, and overall experiences.

Questionnaires.

- **Cognitive Load.** We used the Mental Demand, Physical Demand, and Performance sub-scales of the NASA Task Load Index (NASA-TLX) [33] to measure aspects of presenters' self-assessed cognitive load during the presentation experience. Mental Demand refers to "How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)" [33]. Physical Demand refers to "How much physical activity was required (e.g., pushing, pulling, turning, controlling activating, etc.)" [33]. Performance refers to how successful participants thought they were in accomplishing the goals of the experimental task [33]. We used a 7-point Likert scale for the questions (1 = low workload/poor performance, 7 = high workload/good performance).

- **Group Awareness.** Similar to Schroeder et al. [76] we adapted the descriptive questions from Gutwin and Greenberg's framework [29] to formulate questions about users' perceptions of group awareness. We created questions for the *what* and *where* categories, but we left out the questions about *who* because the participants were only interacting with one other person. We included two questions about artifact awareness from the perspectives of both the participant and their partner. The questions asked participants to rate on a 7-point scale how much they agreed (1 = strongly disagree, 7 = strongly agree). The specific statements are shown in the supplementary materials.

- **Presentation Engagement.** We used a questionnaire created by Webster and Ho [94] measuring overall presentation engagement. We used a 7-point Likert scale for the questions (1 = strongly disagree, 7 = strongly agree). The questions were adapted for both audience and presenter, and they are shown in the supplementary materials.

Video recordings. We recorded videos of the participants in each presentation session and annotated occurrences of the following behaviors:

- **Social interactions.** We annotated each instance of participants gesturing to each other, nodding at each other, smiling at each other, and laughing with each other. We also annotated the beginning and ending timestamps for every instance in which participants explicitly looked at each other.

- **Referencing behaviors.** We annotated each instance of participants establishing references with each other. Based on related studies examining how groups share awareness and

make references in collaborative asymmetric settings [8], we differentiated between the following referencing behaviors: when a participant *pointed* to presentation content, when a participant used *deictic speech* (i.e., spatial language often accompanying pointing such as, "this" or "that"), when a participant used *relative language* (i.e., describing a virtual object in relation to another, e.g., "to the left of the map"), when a participant used a *property reference* (i.e., describing a virtual object based an attribute, e.g., "the orange cylinder"), and when a participant used referential chaining (i.e., describing a virtual object based on a previous reference, e.g., "to the left of the last one").

Interview transcripts. In the semi-structured interview, we asked participants questions in the following categories to provide further insight about the effects of the different presentation modalities:

- **Group Awareness.** We asked participants to describe how they referenced virtual objects and directed their partner's attention during the presentation. We asked them to describe what was straightforward or complicated about making/understanding references and if they could think of any alternative methods they would want to help with this process. We asked presenters whether they experienced differences in their abilities to make/understand references with each presenter modality.

- **Engagement.** We asked participants to describe what was most engaging about the presentation and when they felt most engaged. We also asked audience members whether they felt more engaged with the presentation content or the presenter themselves and how they engaged them. We asked the presenters what techniques they used for engaging the audience with the presentation content and interpersonally and whether they experienced any differences in their abilities to engage the audience with each presenter modality.

- **General Feedback.** We asked participants if they had ideas for additional features they would like to see in this augmented presentation system. We also asked them if they would use this presentation system and to compare it to traditional presentation systems like PowerPoint.

4.3 Analysis

To examine how participants engaged with the presentation, interacted socially, and established group awareness with each presentation system, we aggregated and visualized participant questionnaire responses, annotated behavioral observations in the video recordings of the experimental presentation sessions, and conducted a thematic analysis on the semi-structured interview transcripts.

Questionnaire Responses. We grouped the presenter responses to the cognitive workload questions by presentation modality, computed the medians across sessions, and visualized the results in box plots. For the awareness and engagement questionnaires, we first averaged the scores for all questions per participant to arrive at average awareness and engagement scores per participant per session. Then, we grouped participant responses by role (presenter or audience) and presentation modality before finding their medians and visualizing them in box plots.

---

[3]https://www.census.gov/geographies/reference-files/time-series/geo/centers-population.html

Behavioral Observations. To code participants' behaviors during the study sessions, researcher R1 drafted a codebook of behaviors to observe based on the study's aims of social interaction and group awareness. Then, researchers R1, R2, R3, R4, and R5 met to discuss the code book and suggest changes. The final codebook included codes such as *looking at partner* and *smiling at partner* for social interactions and *pointing to virtual content* and *using deictic speech* for group awareness. Researchers R1 and R2 each coded three presentation videos, while R3, R4, and R5 each coded two sessions using BORIS software [16]. The researchers used the BORIS software [16] to watch synchronized video recordings of the presenter's view, the audience's view, and a third-person view of both participants in the experimental space. Using BORIS, they inserted annotations of the codes at the timestamps at which they were observed. R1 reviewed and validated all coded videos and met with the original coder to resolve disagreements. We grouped participant' social interaction behaviors by role and presentation modality, computed their medians, and visualized them in box plots. We followed the same procedure for the total number of references participants made. We also grouped the data by presentation modality and participant role, and we computed the median number of occurrences of specific reference behaviors per presentation session.

Semi-Structured Interviews. We performed an inductive thematic analysis [5] on interview transcripts to draw common themes and conclusions from our participants' responses to our interview questions. Researchers R1 and R2 independently analyzed the transcripts of the semi-structured interviews to generate codes, then collaboratively reviewed the results and resolved disagreements with all researchers to create the code book. Next, R1 and R2 coded all of the transcripts using the finalized codes. Then, R1 and R2 met to resolve any disagreements. Finally, they formed code categories, organized them into sub-themes, and extracted themes.

4.4 Procedure

When a pair of participants arrived, they first signed a consent form. An experimenter briefed each participant separately on the experimental task and introduced the interface they would use in their assigned experiment condition (e.g., AR HWD for the audience and AR HWD or tablet for the presenter). The presenter was briefed on the presentation content (see subsection 4.1). The presenter was given as much time as they needed to become familiar with the presentation interface and content (typically 5-10 minutes) and informed the experimenter when they were ready to present. The presenter then gave the presentation to the audience. After the presentation, the audience participants filled out the experiment questionnaires, and an experimenter conducted a semi-structured interview. Presenters performed this procedure again with whichever interface they did not use the first time and with a different audience participant. The order in which presenters used the interfaces was counterbalanced between participants. Presenters completed the questionnaires and semi-structured interviews after both presentations were complete.

4.5 Participants

We recruited 18 participants from our organization, 12 as audience and 6 as presenters, for 12 participant pairs. Our audience participants worked in the following roles in our organization: legal (2 participants), finance (2), data science (1), technology/engineering (7). Presenters worked in the following roles: finance (1), technology/engineering (2), human resources (3). We asked participants to report their experience with using AR or VR technology, playing video games, and giving presentations to an in-person audience on a 5-point scale (1 = Never, 2 = Occasionally, 3 = Once a week, 4 = Several times a week, 5 = One or more hours per day). 9 participants had experience with mobile AR. 5 participants had experience with head-worn AR. 7 participants had experience with head-worn VR. For AR and VR experience, participants reported an average of 1.56 (SD = 0.96). For video game experience, participants reported an average of 2.17 (SD = 0.96). For presentation experience, participants reported an average of 2.27 (SD = 0.93). More specifically, 4 presenters reported occasionally giving presentations, and 2 presenters reported giving presentations several times a week. None of our participants knew each other before participating in the experiment.

## 5 Questionnaire and Observational Data

In this section, we first present the responses to the post-study questionnaire and a summary of participant behaviors we observed. Due to the nature of this study as exploratory research intended to inform future research on immersive presentation systems, we report descriptive statistics about our quantitative data without statistical test results [92]. We report only the most relevant results in this section. For a more detailed look at the quantitative results, refer to the supplementary materials.

5.1 Questionnaire Responses

In this section, we report the results of the questionnaires described in section 4.2. All questionnaire responses are on a 7-point Likert scale. These results are visualized in Figure 3.

On the NASA-TLX **cognitive load** questionnaire, presenters reported a slightly higher mental demand score for the symmetric presentation modality (Mdn.: 3.50; IQR: 2.50, 2.25 − 4.75) compared to the asymmetric modality (Mdn.: 3.00; IQR: 2.25, 1.50 − 3.75). Presenters rated their **group awareness** higher in the symmetric presentation modality (Mdn.: 5.75; IQR: 2.03, 4.22 − 6.25) compared to the asymmetric modality (Mdn.: 5.31; IQR: 0.72, 5.00 − 5.72). Similarly, audience participants rated their group awareness higher in the symmetric presentation modality (Mdn.: 6.12; IQR: 0.34, 6.00 − 6.34) compared to the asymmetric modality (Mdn.: 4.62; IQR: 0.38, 4.53 − 4.91). Presenters' **engagement** scores were higher for the symmetric modality (Mdn.: 6.39; IQR: 1.11, 5.75 − 6.86) compared to the asymmetric modality (Mdn.: 6.06; IQR: 1.25, 5.47 − 6.72). Similarly, audience participants' engagement scores were higher for the symmetric modality (Mdn.: 6.50; IQR: 0.53, 6.36 − 6.89) than the asymmetric modality (Mdn.: 5.94; IQR: 0.77, 5.56 − 6.33). Presenters who reported having weekly experience giving presentations tended to rate their ability to engage the audience more highly with a median score of 6.81 (IQR: 0.17, 6.72 − 6.89) compared to 5.61 (IQR: 0.94, 5.25 − 6.19) for those with occasional presentation experience.

(a) Presenter Cognitive Load   (b) Group Awareness   (c) Engagement
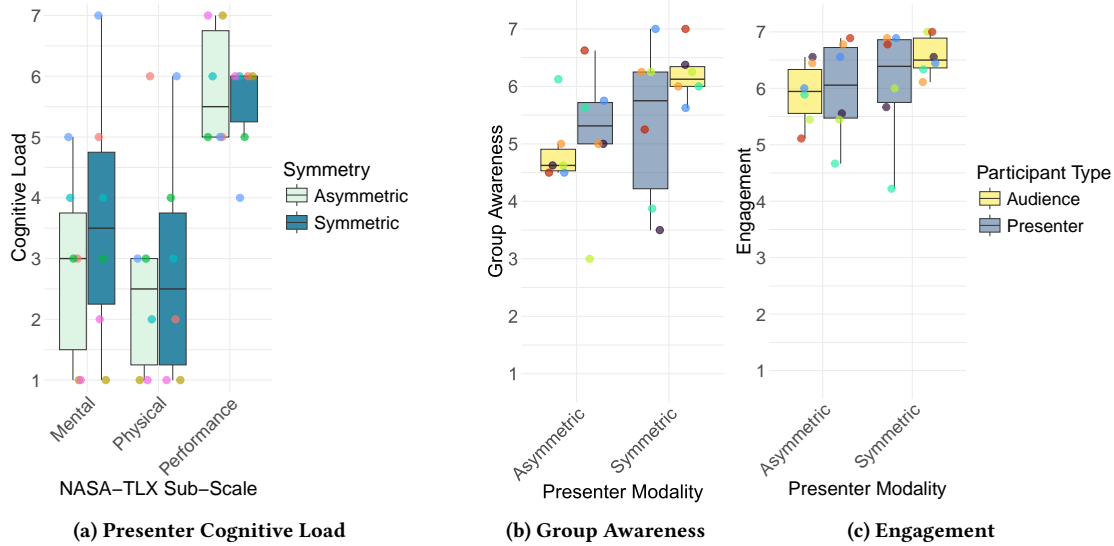
Figure 3: Box plots for questionnaire responses. (a) Average presenter scores for the mental demand (lower is easier), physical demand (lower is easier), and performance (higher is better) categories of the NASA-TLX [33]. (b) Average scores for questions about group awareness. Points represent individual participant responses and are colored by presentation group and presenter: within a presenter modality, points of the same color represent participants in the same presentation session; across presenter modalities, points of the same color represent the same presenter participant. (c) Average scores for questions about the presenter's ability to engage the audience in the presentation and the audience's ratings of their engagement in the presentation. Points represent individual participant responses and are colored by presentation group and presenter.

5.2 Observational Data

In this section, we report quantitative summaries of the behaviors we annotated in the video recordings of the experimental presentation sessions. Presenters exhibited more **social expressions** in the asymmetric modality (Mdn.: 1.50; IQR: 4.50, 0.00 − 4.50) compared to the symmetric modality (Mdn.: 0.50; IQR: 2.50, 0.00 − 2.50). Audience participants also exhibited more social expressions in the asymmetric modality (Mdn.: 4.50; IQR: 13.50, 0.00 − 13.5) compared to the symmetric modality (Mdn.: 2.50; IQR: 1.75, 1.25−3.00). We observed that presenters tended to **look at** the audience for a greater percentage of the presentation when using the asymmetric modality (Mdn.: 6.65%; IQR: 8.30%, 1.65%−9.95%) compared to the symmetric modality (Mdn.: 5.65%; IQR: 10.05%, 1.15% − 11.20%). However, we observed the opposite tendency for audience participants: that they looked at the presenter more when using the symmetric modality (Mdn.: 5.45%; IQR: 8.20%, 3.80%−12.00%) compared to the asymmetric modality (Mdn.: 2.80%; IQR: 4.33%, 1.35% − 5.68%). Participants had more occurrences of **eye contact** in the asymmetric modality (Mdn.: 2.50; IQR: 3.25, 0.50 − 3.75) compared to the symmetric modality (Mdn.: 2.00; IQR: 3.50, 1.00 − 4.50). Similarly, they spent a slightly longer duration of the presentation making eye contact when using the asymmetric modality (Mdn.: 1.00%; IQR: 2.47%, 0.18% − 2.65%) compared to the symmetric modality (Mdn.: 0.38%; IQR: 2.52%, 0.08% − 2.60%).

We observed that presenters tended to make more **references** in the symmetric modality (Mdn.: 11.50; IQR: 13.75, 7.25 − 21.00) compared to the asymmetric modality (Mdn.: 8.50; IQR: 6.80, 4.00 − 10.80). Audience participants made references at a similar frequency

in both symmetric (Mdn.: 5.50; IQR: 1.00, 5.00−6.00) and asymmetric modalities (Mdn.: 7.00; IQR: 5.25, 2.50 − 7.75). We observed that presenters tended to make more *pointing* references when using the symmetric presentation modality (Mdn.: 6.00; IQR: 7.45, 3.75 − 11.20) compared to the asymmetric modality (Mdn.: 3.50; IQR: 6.25, 1.50 − 7.75). Audience members made a similar number of pointing references across both symmetric (Mdn.: 3.00; IQR: 3.50, 1.25 − 4.75) and asymmetric modalities (Mdn.: 4.50; IQR: 4.00, 1.75 − 5.75). Presenters tended to make more *verbal* references when using the symmetric presentation modality (Mdn.: 7.00; IQR: 7.75, 2.00 − 9.75) compared to the asymmetric modality (Mdn.: 2.50; IQR: 4.75, 2.00 − 6.75). Audience members made a similar number of pointing references across both symmetric (Mdn.: 3.00; IQR: 3.50, 0.50 − 4.0) and asymmetric modalities (Mdn.: 2.00; IQR: 2.75, 0.25 − 3.00). Participants mostly used deictic speech for their verbal references.

6 Qualitative Results

In this section, we describe the four themes we derived from our analysis of the interview data: Augmented Presentation Techniques, Spatial Reference and Awareness, Social Interaction and Engagement, Interface and Operation.

6.1 Theme 1: Augmented Presentation Techniques

This theme describes how presenter participants were able to deliver an augmented presentation experience and how audience participants enjoyed specific presentation techniques. It is divided into three categories: Immersive Data Display, Blending 2D and 3D
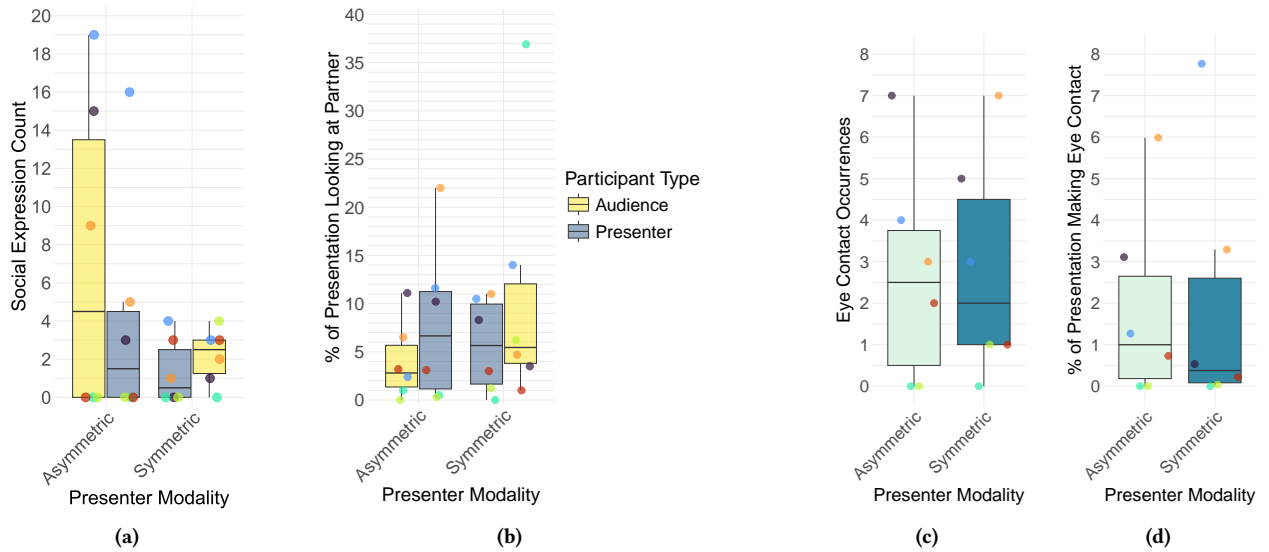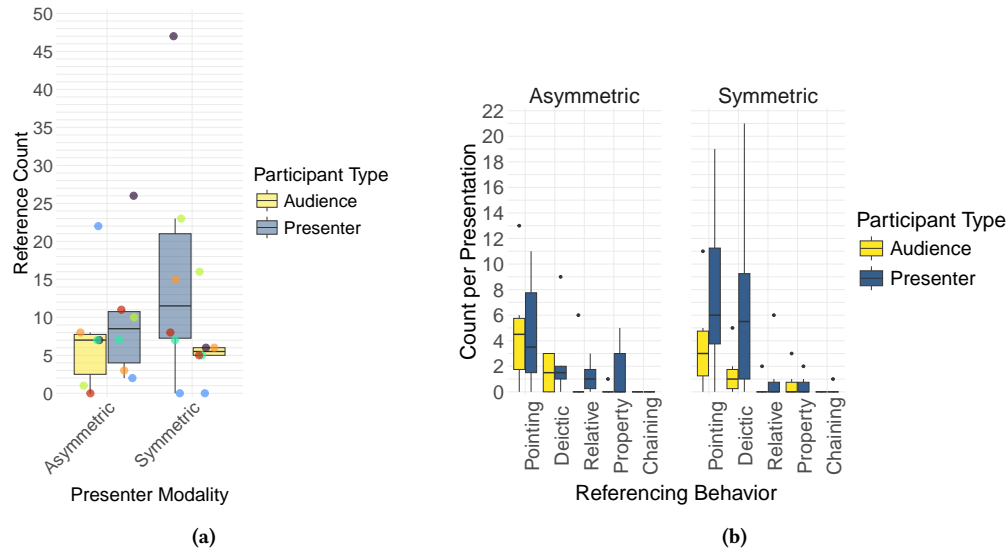
Figure 4: Box plots showing social interactions for each presentation modality. Points are colored by presentation group and presenter: within a presenter modality, points of the same color represent participants in the same presentation session; across presenter modalities, points of the same color represent the same presenter participant. (a) Total number of social expressions (gesturing to each other, nodding at each other, smiling at each other, and laughing) observed for each participant and each presenter modality. (b) Percentage of presentation time participants looked at each other. (c) Number of times participants made eye contact during the presentation. (d) Percentage of the presentation during which participants made eye contact (i.e., looked at each other simultaneously).

Content, and Free-form Exploration. We found that participants, both presenters and audience, were eager to discuss the techniques they would employ when given an augmented presentation system like ours, and many participants expressed a desire to use our system for their daily work.

6.1.1 Immersive Data Display. Our audience participants were particularly impressed by the immersive data displays offered by the augmented presentation system. They appreciated the novel and engaging nature of 3D data visualizations, which gave them fresh perspectives and a deeper understanding of the data compared to plain numbers. The presence and immersion experienced in the 3D environment helped them continuously engage with the presentation content, providing a more focused audience experience. For example, AU02 said, "it was a really cool way to digest information, just to experience something versus having information." Additionally, AU03 said, "not only can I zoom and make [the map data] bigger and then I can also be actually in the map, be actually one with the data, which I found is most intriguing. I was like in the middle of the map [...] this room [could be] filled with data. [I was] fully immersed, and you can actually experience and feel the data." Presenters noted that the immersive data display offered them a more effective way of explaining complex narratives, making information more accessible to comprehend for the audience. PR04 said, "[when] we interact in the 3D space...I think the audience is more interested and engaged." PR06 noted that the dynamic content and the interactivity were very useful for drawing the audience's interest to the presentation.

6.1.2 Blending 2D and 3D Content. Participants pointed out the necessity of having traditional 2D content in an augmented presentation experience, even though 3D visualization may offer a higher engagement than 2D materials. Part of the reason for blending 2D content within a 3D environment is the need to integrate existing work documents into a presentation in a new format. Some participants suggested that they would want to pull a PDF document along with a geo-spatial 3D data visualization, while others mentioned that not every type of data is good to be visualized in 3D. For example, AU09 said that the 3D content could complement existing slideshow materials: "infographics or [...] textual things [are] still easier to put on PowerPoint and [you can show them] in addition to these [3D] data visualizations." AU03 said, "over-utilizing the immersion can be a waste of time depending on the data," and mentioned that it would be beneficial to have some traditional data shown in 2D and conclude with a scene that "brings the whole presentation into one immersive dataset."

6.1.3 Free-form Exploration. Our audience participants enjoyed the opportunity for interactive exploration while the presenter gave the presentation. The interactivity and flexibility allowed them to engage with the content in a more open-ended and personalized manner. Participants valued the freedom and control they had over their presentation experience. Exploring the content independently while the presenter is presenting and talking in parallel enhanced the audience's sense of ownership of the presentation materials. AU01 said, "the presenter could present the slides, [and] I could interact, I could see what I wanted to see[...] Also I have more freedom to look at what I want to, instead of being controlled by the

(a)

(b)

Figure 5: Participant referencing behaviors. (a) Box plot of the total number of references participants made during a presentation for each participant and each presenter modality. Points represent individual participants and are colored by presentation group and presenter: within a presenter modality, points of the same color represent participants in the same presentation session; across presenter modalities, points of the same color represent the same presenter participant. (b) Box plot showing the counts of referencing behaviors observed in a presentation session for each participant type and each presenter modality, including *pointing* to virtual objects, using *deictic* speech (e.g., "this object"), using *relative* language (e.g., "to the left"), using a *property* reference (e.g., "the orange cylinder"), using reference *chaining* (e.g., "next to the last object referenced").

presenter." From the presenter's side, PR05 said, "I'm comfortable letting them explore the 3D object. And I kind of want to stick to my script, and then in a moment I could [...] break into [interacting with the map] if I wanted to [when they ask] can you tell me more about this?" However, this freedom does not come without a cost. Presenter participants pointed out that giving freedom to the audience would require more profound knowledge about the topic, as the questions they would get from the audience can be less predictable. Presenters also expressed concerns during their presentation that their audience looked distracted as they explored the materials independently and engaged excessively with the 3D content.

6.2 Theme 2: Spatial Reference and Awareness

This theme focuses on how participants utilized different spatial reference techniques while giving or taking the augmented presentation and how they tried to gauge their partner's attention. We divide them into four categories: Gesture and Pointing Reference, Verbal Reference, View Sharing and Eye Activity, and Spatial Content Layout.

6.2.1 Gesture and Pointing Reference. Gesture and pointing emerged as the most popular choice for object reference during the augmented presentation experience and was adopted by most presenters and audience participants. Participants frequently used virtual rays, similar to laser pointers in traditional presentations, to direct attention to specific elements in the 3D presentation environment. Participants reported that pointing is natural and easy, and most

believed that their partners were well aware of their pointing gestures. AU06 mentioned, "I used the controller as a laser pointer." While the use of controllers in AR for pointing was found to be straightforward, participants also noted that the accuracy of these gestures could be further improved, especially when the data points visualized in 3D are sometimes dense. Therefore, a few participants also suggested leveraging techniques like eye tracking to enhance control and accuracy further. In addition, some presenters reported that they found it more comfortable to use virtual rays on tablets as the controllers in AR were not as accurate as they expected.

6.2.2 Verbal Reference. While only a few participants adopted verbal communication as their primary method of referencing during presentations, they believed verbal reference to be an effective strategy, primarily when pointing at dense data points, was ineffective. AU06 reported that verbal referencing was effective and helped them when pointing was not working as well as expected. At the same time, verbal references were critical in scenarios when the participants felt that the AR headset limited their visual cues. AU03 reported that verbal communication was necessary "since there is no other indicator to show that I am ready [to be directed] especially since she cannot see my eyes and [...] if we don't verbally communicate that, there is no way for her to know that."

6.2.3 View Sharing and Eye Activity. Presenters often relied on the view-sharing feature provided by our system to ensure that the audience was engaged and focused on the intended aspects of the presentation. Especially when presenters were using tablets as the presenting interface, they monitored where the audience

was looking mainly by the view-sharing window on the tablet. They used this live view to adjust their attention and engagement. PR06 said, "on the tablet I could see directly what the camera is picking up. So you'd be like, oh, their view is wandering, I need to redirect them back to the [visualization]." On the other hand, when both presenters and audience members were wearing AR headsets, presenters tried to use more natural ways to gauge the audience's attention. PR04 said, "I could tell what the audience is looking at by just looking at him [...] if our two worlds are calibrated, then we just look at the same thing [and it] is more natural." PR05 also mentioned that knowing eye activity would help them better understand the audience's interest because "I couldn't read their face at all [behind their headset]."

6.2.4 Spatial Content Layout. Participants desired more structured spatial divisions within the augmented environment, such as dividing the presentation space into multiple quadrants. This organization would help presenters and the audience maintain focus and enhance awareness by clearly delineating different areas for various types of content or interaction. However, presenters also reported challenges related to the expansive nature of AR spaces. Some needed to move their heads frequently between viewing presentation content, speaker notes, and the audience's faces. This constant movement made maintaining continuous focus and awareness of the audience's attention difficult, potentially impacting their overall understanding of the partner's engagement. In contrast, the tablet interface was noted for its more integrated components, allowing presenters to switch their visual attention quickly and effortlessly between different elements. This integration provided a higher level of awareness and ease of use, enabling presenters to stay more attuned to the audience's reactions and attention without the physical strain posed by AR environments.

6.3 Theme 3: Social Interaction and Engagement

This theme relates to participants being able to socially interact and engage with each other during an augmented presentation experience. We have four categories for this theme: Engagement Techniques and Judgment, Social Co-presence, Eye Contact for Social Connection, and Dialogue Encouragement.

6.3.1 Engagement Techniques and Judgment. During the presentation experience, presenters employed various techniques to engage the audience and judge their level of focus and involvement. Presenters expressed concern that audience members seemed too focused on the immersive 3D content, often at the cost of maintaining visual contact with the presenter. To counter this, presenters frequently used verbal engagement strategies, such as asking questions, to encourage audience interaction. Additionally, some presenters used 3D content manipulation, such as panning and zooming in and out on maps, to guide the audience's attention and maintain engagement throughout different presentation sections.

To gauge audience involvement, presenters observed how the audience interacted with the 3D content and noted the types of questions asked during the presentation. PR06 relied on the live audience view in the tablet for this purpose and said, "it was easier to tell [the audience's engagement level] on the tablet because I

could see their like vision moving around as like things were changing." PR03 mentioned that the familiarity of the tablet interface enabled them to engage the audience better: "Because I was more comfortable with the tablet, I think I was a lot more engaging in that presentation. I didn't really have to think too much about what I was doing." On the other hand, most of the audience appreciated the presenters' attempt to engage them with verbal cues and thought that having someone talk to them and walk them through the presentation materials by voice helped them better understand and experience them. For example, AU08 said, "Just having a voice-over to navigate the experience was really good. That definitely made it engaging." AU04 mentioned the need for more body and visual cues by virtual avatars, with suggestions for stronger social connections to help bridge the engagement gap.

6.3.2 Social Co-presence. Participants using AR headsets reported a stronger sense of intimacy and connection, with interaction styles—such as gestures and face-to-face conversations—feeling more natural than those using tablets. The symmetric setup, where both the presenter and audience used headsets, made participants feel as though they were together, experiencing the presentation content simultaneously. For example, AU02 noted, "it made me feel like we were more kind of experiencing something together, 'cause we had the same equipment on versus just kind of the presenter just being like a separate person over there just showing the slides. It did make me feel like we are experiencing together." In contrast, the asymmetric setup positioned the presenter as more of a solo guide. Tablets gave the presenter a greater sense of anonymity and less direct control or involvement with the virtual content in the space. AU07 said, "from my perspective, I thought I was in my own environment [...] I didn't really see him in the world. I just thought of him as a voice over until started seeing the [virtual pointing ray]." Some presenters also expressed uncertainty about whether the audience was fully engaged with the presentation or simply exploring independently, leading to concerns that the audience members might have been disconnected and immersed in their world. For example, PR05 said, "It made me excited that [the audience] could do their own exploration. And I was like merely their guide rather than like presenting it to them [...] but I couldn't pick up on many cues as to whether they might have been disinterested or distracted. I can't see their face."

6.3.3 Eye Contact for Social Connection. In the interviews, participants consistently highlighted the importance of eye contact for maintaining social connection during AR presentations, but many encountered challenges due to the headset's limitations. AU03 mentioned that they tried to make eye contact with their partner but stopped, feeling that their partner likely couldn't see their eyes, leading to a sense of uni-directional interaction. Others reported that the headset restricted their peripheral view, while some had to step back to see the entire desk or map, further complicating attempts at social engagement. AU11 and PR03 noted that "the headset was a bit of burden", making it difficult for the audience to rotate their heads or look at their partners naturally. In the asymmetric setup, presenters using tablets were better able to observe the audience's body language, and it was easier to socially interact with the audience when using the tablet, as "the headset feels like a physical blocker," hindering direct engagement compared to the

tablet. AU03 noted, "oh she can't see my eyes now. I can tell when she was looking at me, she was a good presenter. And [...] she knows that I am looking at her now, but it's a weird masked feeling where you know she is not able to tell if I have my eyes closed, and she would have no idea."

6.3.4 Dialogue Encouragement. The participants reported that the augmented presentation offered greater visual and social dynamism, leading to improved engagement between the audience and presenter and encouraging more interaction and Q&A. Participants, such as AU01 and PR05, observed that the enhanced interactivity naturally prompted more feedback and questions from the audience. However, both presenters and audience members noted that this type of presentation demands more time for discussion, emphasizing that presenters need to be well-prepared for unpredictable questions. For example, PR05 said, "In an ideal scenario, I would go in the immersive environment with [the audience] and we'd have a conversation about [the presentation content]. I'd be knowledgeable enough about it that we would just have a conversation." Additionally, AU01 said, "the presenter has to be more creative in giving the narrative [...] because the [audience] can do anything [...] You have to be more knowledgeable about what you are going to present because there will be more types of questions that come out." This requires a higher familiarity with the material than traditional PowerPoint presentations. Participants felt that the interactivity of augmented presentations fosters more conversation, especially for small focus groups, with AU05 adding that whereas in "some mundane [presentations], people start to fall asleep," but the immersive presentation "keeps you up."

Some presenters also commented on the differences in the interface. The tablet was generally seen as a familiar and easy-to-use tool with less overhead for the presenters that encouraged social interactivity. For example, one presenter felt that the headset didn't provide the level of control they expected, making them concerned about unintentionally "jarring" the audience. Only one presenter expressed that the headset allowed more natural conversation as the tablet does not allow enough space for gesture cues.

6.4 Theme 4: Interface and Operation

This theme presents participants' feedback on the interface impact of AR and tablet devices and how different interfaces affected their operation while conducting the presentation. This theme has three categories: Control Competition, Perception of Interfaces, and Technology Adaptation.

6.4.1 Control Competition. Participants reported managing shared control and ownership over the 3D map between the presenter and the audience. Several participants reported issues with competing for power, resulting in jumpy and disjointed map manipulations. PR01 specifically pointed out that this competition hindered the fluidity of the presentation. To avoid such conflicts, some audience members opted to avoid directly interacting with the map, instead pointing to areas of interest to guide their partners subtly. Others suggested more structured solutions, such as implementing a turn-based control system, allowing the audience to take some extra time to freely explore the content after each slide, especially since the immersive and interactive data display requires more time to

comprehend, as mentioned by AU03: "every time the data changed, it took time to talk through it [...] If you had a PowerPoint presentation, that would be one thing, but if you change the map and the landscape of the environment you are in, it's like you are having to [...] take back in the changed environment." Some participants, including AU01 and PR01, proposed giving the audience a dedicated window or interface for controlling the map to reduce competition. PR04 also expressed the need for the ability to reset the map after the audience had manipulated it to maintain the presenter's control over the presentation flow.

6.4.2 Perception of Interfaces. Participants reported that different setups could vary the presenter's and audience's experience during an augmented presentation. Three presenters preferred the tablet because it provided a familiar interface, a better overview, and easier presentation control. For example, PR06 said that with the tablet, "it was easier to keep track of everything" whereas the headset interface was "a bit distracting". Furthermore, R03 and PR06 noted that positioning the presentation controls in the headset interface caused additional overhead. At the same time, the headset was perceived as fostering a more collaborative and natural interaction due to its symmetric nature, where both the presenter and the audience engage with the content equally. The tablet, on the other hand, reinforced the role of the presenter, with participants feeling more in control and in a traditional presentation mode while holding it. AU03 added a perspective on the context of co-located interactions, expressing a preference and necessity for maintaining a connection via the physical environment. They argued that just as it doesn't make sense to have Zoom meetings when both users are physically present; using excessive immersive space in situations where users are co-located also can make people feel unnecessary.

6.4.3 Technology Adaptation. Our participants expressed strong interest in using the system for their daily work, but they also highlighted several factors necessary for smoother adoption. One key aspect was the need for an easy-to-use authoring method, as creating 3D visualizations for presentations requires both time and technical skills. Both presenters and audience members mentioned the importance of having simple tools to quickly build 3D presentations and easily map and convert 2D data into 3D formats. Participants also emphasized the value of standardized design patterns, templates, or frameworks for 3D scene slides that anyone could use readily, reducing the barrier to preparing such an experience.

Some participants also reported that working in 3D was cognitively demanding. Presenters, in particular, found it challenging to manage speaker notes, manipulate 3D content, and engage the audience simultaneously. PR05 noted, "Being immersed was overwhelming, so it was harder to focus on my story and my speaker notes." Participants also wanted physical interaction within the AR headset to feel more natural and for the device to be more comfortable. The frequent head movement required in an immersive mode made them desire lighter headsets for better comfort. A few participants mentioned feeling dizzy or experiencing headaches during the presentation, underscoring the importance of improved hardware for better user adoption and overall experience.

7 Discussion

This section presents design guidelines for creating immersive presentation systems derived from our main findings. Additionally, we outline the limitations of the present study. Lastly, we present ideas for future work to build on our findings.

7.1 Design Guidelines for Immersive Presentation Systems

This section offers design guidelines for researchers and practitioners aiming to create co-located immersive presentation systems that maximize user engagement, with insights on selecting suitable presentation formats and key considerations for content preparation to ensure meaningful and interactive experiences.

7.1.1 Support balanced interactivity and audience exploration. In the interviews, participants reported feeling engaged with the novel presentation and being particularly drawn to the content's immersiveness and interactivity. They also noted how the 3D content complements more traditional 2D presentation content. To best support interactivity in augmented presentations, it is important to use data visualization that is well-suited for 3D display, as there are significant trade-offs between the benefits of 3D interactivity and the challenges it presents [36, 46, 72]. On the positive side, 3D interactivity allows the audience to explore the content independently, promoting a deeper understanding of the subject matter [46] and potentially encouraging more conversation and Q&A, as observed in our study. However, as our participants noted, this increased audience engagement also places more demands on the presenter, who must be highly familiar with the details of the dataset. The presenter needs to be prepared for unpredictable audience exploration [46] and questions or observations from the audience, and potentially some difficulty controlling the pace of the presentation. This can shift the presentation toward a more collaborative experience, with the audience influencing how the content is displayed or introducing new topics based on their discoveries. As indicated by our qualitative data results, this sense of collaboration is especially stronger in a symmetric setup, characterized by a sense of equality, which fosters a stronger role for the audience in terms of their control capability. We also observe a similar trend in the quantitative data, suggesting that a symmetric setup could potentially lead to better group awareness. This aligns with the findings by Drey et al. that a symmetric setup improves communication and workspace awareness in paired VR learning [13]. However, our presentation scenario is a more uni-directional and structured experience compared to paired learning. The shift from guided presentation towards collaboration can reduce the presenter's control and ownership over the flow of the presentation. This level of audience interactivity needs to be carefully considered depending on the style and purpose of the presentation. Different from Drey et al.'s guideline that the experience should always allow both users to move and interact freely in the VE and that symmetrical setups is always superior, our work highlights the unique trade-offs between symmetric and asymmetric settings in augmented presentation. **We recommend balancing interactivity with the presenter's ability to guide and manage the overall experience based on the presentation goals, recognizing that symmetric setups can**

potentially facilitate easy collaboration with the audience, whereas asymmetric setups offer the presenter good control and direction.

7.1.2 Support social interaction between the presenter and audience. In observing participants' behaviors during the presentations, there were more social interaction occurrences (e.g., nodding or gesturing to each other, smiling) when presenters used the asymmetric presenter modality compared to the symmetric modality. Both presenters and audience also made more eye contact with greater frequency and for longer durations when using the asymmetric presentation modality. During interviews, participants noted that the reduced observable social cues caused by wearing the HWD led to lower-quality social interaction. This finding aligns with related research suggesting that XR headsets may limit typical social interactions because they occlude a large portion of the wearer's face, therefore preventing natural eye contact and complete observation of facial expressions [53, 90].

On the other hand, we observed that audience members tended to spend more time looking at the presenter when they used the symmetric modality. Additionally, in interviews, audience members reported feeling like they were in a separate world from the presenter when they used the asymmetric presentation modality because the presenter was not immersed in the same AR space. In the symmetric condition, participants reported the opposite: they felt part of the same experience. While we did not ask participants about their perceptions of co-presence [31], future work should investigate this presentation scenario further in this regard.

It is worth considering design compromises to resolve the tension between reduced social interactions and increased time spent looking at the presenter in the symmetric scenario. One solution explored in the literature is to display an HWD user's eyes on the outside of the display to restore eye gaze cues and awareness and increase social presence with others [3, 7, 51, 52]. The development of photorealistic avatars [9, 86] and enhanced facial animation generation [85, 100] in addition to Apple's Personas[4] and Meta's codec avatars[5] could be used to overlay an avatar on users with HWDs to preserve observation of their social signals. From the presenter's side, when using the tablet interface it may be useful to include an AR view so that the presenter may peer into and interact with the AR space directly. Another option is for the audience to use a tablet with an AR view, which will place both the presenter and audience in the same world and reduce social interaction blockers for both users. However, we recommend that the audience use the AR HWD due to its advantages in permitting more direct interaction and increased immersiveness.

In cases that require emphasizing social connection between the presenter and audience through typical social behaviors, we recommend reducing equipment that inhibits the presenter's social expressions, e.g., by using a tablet to control the presentation. When it is more important that the presentation feel like a shared experience, we recommend that both the presenter and audience use AR HWDs.

---

[4]https://www.apple.com/apple-vision-pro/
[5]https://about.meta.com/realitylabs/codecavatars/

7.1.3 Support intuitive referencing and awareness. As noted in our observations of participant behavior and their interviews, participants made spatial references to virtual content using both pointing and verbal communication during the presentation to direct their partner's attention. As shown in Figure 5a, presenters tended to make more references when using the symmetric modality. More specifically, as shown in Figure 5b, when using the symmetric presentation modality, presenters pointed and used deictic phrases to refer to content more frequently than in the asymmetric modality. Additionally, as shown in Figure 3b, audience members appeared to experience greater group awareness with the presenter (e.g., they could better understand each other's references and attention) in the symmetric condition. In the interviews, participants agreed that referencing was intuitive in the symmetric modality. Still, the AR HWD blocking gaze direction prevented complete awareness of where their partner's attention was and what they were interested in. This point is another reason to consider displaying the HWD user's eyes outside of the headset [3, 7, 51, 52] or adding visualizations to represent a user's orientation and gaze [15, 63].

Some participants reported successfully using the tablet's view sharing and pointing ray features to make references, which aligns with related research employing similar referencing techniques in asymmetric collaboration scenarios [38, 76]. However, the lower scores for the group awareness questionnaire, the fewer references made, and participants' comments about feeling as though they are in separate worlds for the asymmetric modality suggest these features are not equivalent to a symmetric experience in providing a common ground for spatial referencing. This highlights a unique challenge in AR compared to the VR setting studied by Drey et al. [13] — the prominent presence of the physical environment in AR for both co-located parties necessitates more aligned levels of awareness and referential cues, for which the less direct signaling approaches could fall short. One participant suggested the tablet could have an AR view option to address this issue. **We recommend providing view and interaction symmetry (e.g., by both presenter and audience using AR HWDs or by providing an AR view through the presenter's tablet) in cases where group awareness is critical, such as when the presentation calls for lots of shared attention and interaction around the presentation content.**

7.1.4 Consider interface familiarity and ergonomics. Immersive technology offers a much larger space to display content and provides a more flexible presentation control interface, allowing for greater expressiveness in presentations. However, this increased space and flexibility can introduce challenges, such as interaction and visual fatigue, disorientation, and higher cognitive and physical effort. Our quantitative data suggests a trend toward lower mental demand when using the tablet interface, and in the interviews, some participants reported having difficulty interacting with the AR HWD interface. While presenters ideally would adjust the interface to suit their ergonomic needs, in practice, presenters are often under heavy cognitive load during a presentation. They have to multitask for a presentation, such as engaging the audience, delivering expressive narration, and monitoring audience behavior, making it difficult to focus on arranging an ergonomic interface layout based on the changing position of the audience.

A key finding from the interviews was that three of the presenters preferred using the tablet over the AR HWD because it provided better integration of content and control, allowing them to manage everything — control the slides, observe the content and the audience's focus, read scripts, and engage with the audience — within one centralized space, requiring far less physical movement. Presenters also commented that the familiarity of the tablet interface with other presentation tools they already use made it easier to deliver a good presentation. This familiarity integration becomes even more critical in real-world workplace settings, where presenters must deliver multiple presentations in succession. **We recommend using highly integrated and conventional interfaces to minimize mental load and make the presentation system more practical for sustained use, particularly in professional environments where efficiency and endurance are vital.** In contrast, a sparse and disjointed interface layout can quickly become labor-intensive and lead to faster fatigue, undermining the benefits of immersive presentation technology.

7.1.5 Pre-presentation Preparation and Real-time Content Manipulation. When designing authoring systems for augmented presentations, it is important to consider both pre-presentation preparation and real-time manipulation during the presentation itself. Pre-presentation authoring should resemble the process of preparing traditional slides, where the presenter maintains complete control over the content, structure, and flow ahead of time. This method helps to ensure a smooth delivery. The advantage is that the presenter can focus on delivering the content without manipulating the data or visuals in real-time, reducing cognitive load during the presentation.

In the interviews, our participants highlighted that having tools and design templates for easy creation of augmented presentations, such as 2D to 3D data mapping tools and PowerPoint slide conversion into 3D formats, would greatly ease the onboarding process. This approach would allow us to build on established user-centered design principles from web and desktop interfaces [10]. At the same time, our participants indicated a preference for using 2D and 3D materials for different situations and purposes. Traditional documents like PDFs and existing 2D slides can serve well without introducing additional preparation overhead, while 3D visual content is ideal for interactive data exploration and sense-making. Especially when it comes to 3D content experience, real-time manipulation of presentation content allows the presenter to tailor the experience to the specific needs and interests of the audience. As we derived from the interviews, this dynamic approach can enhance audience engagement by making the presentation more interactive and personalized. However, it also introduces a higher cognitive burden for the presenter, as they must manage both the content and its presentation on the fly while engaging the audience. This type of authoring requires a balance between flexibility and control, ensuring that while the presenter can make real-time adjustments, the interface and tools do not overwhelm them with complexity. **We recommend designing a system that supports intuitive pre-planning and real-time adjustments, effectively integrating 2D and 3D content based on actual needs of interactivity to enhance the effectiveness and versatility of augmented presentations. By doing so, presenters can leverage the strengths**

of both content types to deliver more compelling and contextually relevant presentations while minimizing friction during the preparation process.

7.2 Limitations

One limitation of our exploratory user study is that it involved a small sample size (n = 12 groups). Our data showed variance across groups, potentially due to individual participant differences (e.g., presenters who minimally engaged their audiences). While our results still provided valuable insights, the limited number of participants may impact the generalizability of the findings to the broader population.

In real-world presentation scenarios, presenters are often familiar with their presentation material before the presentation. However, in our user study, participants were not experts on the subject matter and had limited time to familiarize themselves with the presenter interfaces and practice the presentation, which limits the generalizability of our findings. While no participants complained about having enough time to practice, this could have reduced presenters' confidence in the system and the presentation content, which could have also reduced their performance or social interactions. A lack of confidence in the system or content might explain why some presenters hardly interacted with their audience during the presentation and focused mainly on advancing the slides and communicating the messages on the speaker notes.

7.3 Future Work

To address the lack of social behaviors exhibited by some presenters, one avenue for future work is to provide automatic and explicit cues to the presenter to engage the audience, e.g., by periodically reminding them to interact socially with the audience. Additionally, a system could detect the audience's engagement level throughout the presentation (e.g., [35]) and automatically provide feedback to the presenter about which kinds of content or interactions are most engaging for the audience.

Likewise, we observed that the ability of the audience to explore the data on their own was seen as a benefit by both audiences and presenters, with the latter sometimes encouraging this exploration. However, there will still be times when the presenter demands the audience's full attention, such as when introducing an important concept. Thus, proper mechanisms in facilitating transitions between controlled and exploration presentation phases may be investigated in the future, whether it is to provide explicit (or implicit) cues for when the audience is expected to explore the data or even to devise appropriate techniques to "gracefully interrupt" an audience member who may be engrossed in the data.

While we focus primarily on an intimate tabletop setting for this work, there are myriad more scenarios and contexts where presentations are often conducted, which can benefit from AR. The number of audience members can easily vary, ranging from small groups (e.g., a family) to large crowds (e.g., a conference talk), and there can be more than one presenter, be it multiple presenters swapping once throughout the presentation or them interweaving their narration between each other. These simple differences introduce new challenges, for example, how larger groups of people share the physical space during co-located augmented presentations, how presenter(s)

keep awareness of what their audience member(s) are looking at, and how the audience member(s) might want to explore the data for themselves. Future work could also explore how using other virtual augmentations such as visualizing the AR HWD user's eyes on the outside of the headset (e.g., as studied in research [3, 7, 51, 52] and implemented in the Apple Vision Pro) or using VR in co-located presentations affect users' experiences.

Additionally, our study results may be influenced by novelty effects, as some participants were new to AR technology. A future longitudinal study could help us better understand how users' social interactions and engagement might evolve as they become more familiar with the system. Such a study could also involve usage in realistic scenarios, where participants present their own materials, allowing for a more comprehensive evaluation of the system's usability and effectiveness over time and in real-world conditions.

The space that the augmented presentation is occurring in can also vary, mainly when the presentation is situated around some physical context [96]. For example, a presentation for prospective home buyers might involve walking around different parts of the house, or a presentation for car dealerships might involve buyers physically inspecting multiple cars side by side, guided by a presenter. The need to navigate and coordinate presentation content in large spaces poses new challenges, for example, how the presenter might be aware of the location of presentation content without (or even with) an AR HWD, and how they might ensure proper attention guidance and engagement for audiences who may become distracted in a more unconstrained space.

8 Conclusion

Motivated by the use of AR to facilitate immersive and engaging presentations to audiences, we developed an augmented presentation prototype system and investigated how the choice of presenter modality—immersive AR (i.e., symmetric) versus non-immersive tablet (i.e., asymmetric)—influences engagement, group awareness, and social interaction between the presenter and audience. We conducted a user study of 12 participant pairs wherein a presenter, either with an AR HWD or a tablet, delivered a presentation to an audience member wearing an AR HWD. Based on participants' experiences, we derived four themes around presentation techniques, spatial reference, social interaction, and interface impact. We provide design guidelines for immersive presentation systems based on these findings and propose several new directions to support future research and real-world applications of augmented presentations.

References

[1] John N Bassili. 1979. Emotion recognition: the role of facial movement and the relative importance of upper and lower areas of the face. *Journal of personality and social psychology* 37, 11 (1979), 2049. https://doi.org/10.1037/0022-3514.37.11.2049
[2] Verena Biener, Snehanjali Kalamkar, Negar Nouri, Eyal Ofek, Michel Pahud, John J Dudley, Jinghui Hu, Per Ola Kristensson, Maheshya Weerasinghe, Klen Čopič Pucihar, et al. 2022. Quantifying the effects of working in VR for one week. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3810–3820. https://doi.org/10.1109/TVCG.2022.3203103
[3] Evren Bozgeyikli, Lal "Lila" Bozgeyikli, and Victor Gomes. 2024. Googly Eyes: Exploring Effects of Displaying User's Eyes Outward on a Virtual Reality Head-Mounted Display on User Experience. In *2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)*. 979–989. https://doi.org/10.1109/VR58804.2024.00117
[4] Erin Bradner and Gloria Mark. 2002. Why Distance Matters: Effects on Cooperation, Persuasion and Deception. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work* (New Orleans, Louisiana, USA)

(CSCW '02). Association for Computing Machinery, New York, NY, USA, 226–235. https://doi.org/10.1145/587078.587110

[5] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis.* American Psychological Association. https://doi.org/10.1080/17439760.2016.1262613

[6] Davide Calandra, Filippo Gabriele Prattico, Massimo Migliorini, Vittorio Verda, and F. Lamberti. 2021. A Multi-role, Multi-user, Multi-technology Virtual Reality-based Road Tunnel Fire Simulator for Training Purposes. 96–105. https://doi.org/10.5220/0010319400960105

[7] Liwei Chan and Kouta Minamizawa. 2017. FrontFace: facilitating communication between HMD users and outsiders using front-facing-screen HMDs. In *Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Vienna, Austria) *(MobileHCI '17).* Association for Computing Machinery, New York, NY, USA, Article 22, 5 pages. https://doi.org/10.1145/3098279.3098548

[8] Jeffrey W. Chastine, Kristine Nagel, Ying Zhu, and Luca Yearsovich. 2007. Understanding the design space of referencing in collaborative augmented reality environments. In *Proceedings of Graphics Interface 2007* (Montreal, Canada) *(GI '07).* Association for Computing Machinery, New York, NY, USA, 207–214. https://doi.org/10.1145/1268517.1268552

[9] Jianchun Chen, Jian Wang, Yinda Zhang, Rohit Pandey, Thabo Beeler, Marc Habermann, and Christian Theobalt. 2024. EgoAvatar: Egocentric View-Driven and Photorealistic Full-body Avatars. In *SIGGRAPH Asia 2024 Conference Papers (SA '24).* Association for Computing Machinery, New York, NY, USA, Article 102, 11 pages. https://doi.org/10.1145/3680528.3687631

[10] Mengyu Chen, Andrés Monroy-Hernández, and Misha Sra. 2021. SceneAR: Scene-based Micro Narratives for Sharing and Remixing in Augmented Reality. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* 294–303. https://doi.org/10.1109/ISMAR52148.2021.00045

[11] 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 (April 2022), 27 pages. https://doi.org/10.1145/3512909

[12] Arindam Dey, Thammathip Piumsomboon, Youngho Lee, and Mark Billinghurst. 2017. Effects of Sharing Physiological States of Players in a Collaborative Virtual Reality Gameplay. 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, 4045–4056. https://doi.org/10.1145/3025453.3026028

[13] 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

[14] Steve Dübel, Martin Röhlig, Heidrun Schumann, and Matthias Trapp. 2014. 2D and 3D presentation of spatial data: A systematic review. In *2014 IEEE VIS International Workshop on 3DVis (3DVis).* 11–18. https://doi.org/10.1109/3DVis.2014.7160094

[15] Austin Erickson, Nahal Norouzi, Kangsoo Kim, Ryan Schubert, Jonathan Jules, Joseph J LaViola Jr, Gerd Bruder, and Gregory F Welch. 2020. Sharing gaze rays for visual target identification tasks in collaborative augmented reality. *Journal on Multimodal User Interfaces* 14, 4 (2020), 353–371. https://doi.org/10.1007/s12193-020-00330-2

[16] Olivier Friard and Marco Gamba. 2016. BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in Ecology and Evolution* 7, 11 (2016), 1325–1330. https://doi.org/10.1111/2041-210X.12584

[17] Susan R. Fussell, Robert E. Kraut, and Jane Siegel. 2000. Coordination of communication: effects of shared visual context on collaborative work. In *Proceedings of the 2000 ACM Conference on Computer Supported Cooperative Work* (Philadelphia, Pennsylvania, USA) *(CSCW '00).* Association for Computing Machinery, New York, NY, USA, 21–30. https://doi.org/10.1145/358916.358947

[18] Susan R. Fussell, Leslie D. Setlock, and Robert E. Kraut. 2003. Effects of head-mounted and scene-oriented video systems on remote collaboration on physical tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03).* Association for Computing Machinery, New York, NY, USA, 513–520. https://doi.org/10.1145/642611.642701

[19] Maia Garau, Mel Slater, Simon Bee, and Martina Angela Sasse. 2001. The impact of eye gaze on communication using humanoid avatars. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Seattle, Washington, USA) *(CHI '01).* Association for Computing Machinery, New York, NY, USA, 309–316. https://doi.org/10.1145/365024.365121

[20] Tiago Gonçalves, Ana Paula Afonso, and Bruno Martins. 2016. Why not both?–combining 2D maps and 3D space-time cubes for human trajectory data visualization. In *Proceedings of the 30th International BCS Human Computer Interaction Conference.* BCS Learning & Development. https://doi.org/10.14236/ewic/HCI2016.22

[21] Matt Gottsacker, Mengyu Chen, David Saffo, Feiyu Lu, and Blair MacIntyre. 2023. Hybrid User Interface for Audience Feedback Guided Asymmetric Immersive Presentation of Financial Data. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct).* 199–204. https://doi.org/10.1109/ISMAR-Adjunct60411.2023.00046

[22] Matt Gottsacker, Nahal Norouzi, Kangsoo Kim, Gerd Bruder, and Greg Welch. 2021. Diegetic Representations for Seamless Cross-Reality Interruptions. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR).* 310–319. https://doi.org/10.1109/ISMAR52148.2021.00047

[23] Matt Gottsacker, Raiffa Syamil, Pamela Wisniewski, Gerd Bruder, Carolina Cruz-Neira, and Gregory Welch. 2022. Exploring Cues and Signaling to Improve Cross-Reality Interruptions. In *2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct).* 827–832. https://doi.org/10.1109/ISMAR-Adjunct57072.2022.00179

[24] 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

[25] Luke E Grzeskowiak, Alicia E Thomas, Josephine To, Emily Reeve, and Adam J Phillips. 2015. Enhancing continuing education activities using audience response systems: a single-blind controlled trial. *Journal of Continuing Education in the Health Professions* 35, 1 (2015), 38–45. https://doi.org/10.1002/chp.21269

[26] 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

[27] Pavel Gurevich, Joel Lanir, and Benjamin Cohen. 2015. Design and Implementation of TeleAdvisor: a Projection-Based Augmented Reality System for Remote Collaboration. *Computer Supported Cooperative Work (CSCW)* 24 (2015), 527 – 562. https://doi.org/10.1007/s10606-015-9232-7

[28] Pavel Gurevich, Joel Lanir, Benjamin Cohen, and Ran Stone. 2012. TeleAdvisor: A Versatile Augmented Reality Tool for Remote Assistance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12).* Association for Computing Machinery, New York, NY, USA, 619–622. https://doi.org/10.1145/2207676.2207763

[29] 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. https://doi.org/10.1023/A:1021271517844

[30] Brian D. Hall, Lyn Bartram, and Matthew Brehmer. 2022. Augmented Chironomia for Presenting Data to Remote Audiences. In *Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology* (Bend, OR, USA) *(UIST '22).* Association for Computing Machinery, New York, NY, USA, Article 18, 14 pages. https://doi.org/10.1145/3526113.3545614

[31] Chad Harms and Frank Biocca. 2004. Internal consistency and reliability of the networked minds measure of social presence. In *Seventh annual international workshop: Presence,* Vol. 2004. Universidad Politecnica de Valencia Valencia.

[32] John Harris and Mark Hancock. 2019. To Asymmetry and Beyond! Improving Social Connectedness by Increasing Designed Interdependence in Cooperative Play. 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, 1–12. https://doi.org/10.1145/3290605.3300239

[33] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[34] Faradillah Haryani and N Ayuningtyas. 2021. The impact of interactive online learning by Pear Deck during COVID-19 pandemic era. In *Journal of Physics: Conference Series,* Vol. 1957. IOP Publishing, 012006. https://doi.org/10.1088/1742-6596/1957/1/012006

[35] Mariam Hassib, Stefan Schneegass, Philipp Eiglsperger, Niels Henze, Albrecht Schmidt, and Florian Alt. 2017. EngageMeter: A System for Implicit Audience Engagement Sensing Using Electroencephalography. 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, 5114–5119. https://doi.org/10.1145/3025453.3025669

[36] Petra Isenberg, Bongshin Lee, Huamin Qu, and Maxime Cordeil. 2018. *Immersive Visual Data Stories.* Springer International Publishing, Cham, 165–184. https://doi.org/10.1007/978-3-030-01388-2_6

[37] Pascal Jansen, Fabian Fischbach, Jan Gugenheimer, Evgeny Stemasov, Julian Frommel, and Enrico Rukzio. 2020. ShARe: Enabling Co-Located Asymmetric Multi-User Interaction for Augmented Reality Head-Mounted Displays. 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, 459–471. https://doi.org/10.1145/3379337.3415843

[38] Janet G Johnson, Danilo Gasques, Tommy Sharkey, Evan Schmitz, and Nadir Weibel. 2021. Do You Really Need to Know Where "That" Is? Enhancing Support for Referencing in Collaborative Mixed Reality Environments. 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 514, 14 pages. https://doi.org/10.1145/3411764.3445246

[39] Jill King and Julie Markant. 2020. Individual differences in selective attention and scanning dynamics influence children's learning from relevant non-targets in a visual search task. Journal of Experimental Child Psychology 193 (2020), 104797. https://doi.org/10.1016/j.jecp.2019.104797

[40] Robert E. Kraut, Susan R. Fussell, and Jane Siegel. 2003. Visual Information as a Conversational Resource in Collaborative Physical Tasks. Human–Computer Interaction 18, 1-2 (2003), 13–49. https://doi.org/10.1207/S15327051HCI1812_2

[41] Yoshiki Kudo, Anthony Tang, Kazuyuki Fujita, Isamu Endo, Kazuki Takashima, and Yoshifumi Kitamura. 2021. Towards Balancing VR Immersion and Bystander Awareness. Proc. ACM Hum.-Comput. Interact. 5, ISS, Article 484 (nov 2021), 22 pages. https://doi.org/10.1145/3486950

[42] Mohammad Amin Kuhail, Areej ElSayary, Shahbano Farooq, and Ahlam Alghamdi. 2022. Exploring Immersive Learning Experiences: A Survey. Informatics 9, 4 (2022). https://doi.org/10.3390/informatics9040075

[43] T. Kurata, N. Sakata, M. Kourogi, H. Kuzuoka, and M. Billinghurst. 2004. Remote collaboration using a shoulder-worn active camera/laser. In Eighth International Symposium on Wearable Computers, Vol. 1. 62–69. https://doi.org/10.1109/ISWC.2004.37

[44] Sergej Lackmann, Pierre-Majorique Léger, Patrick Charland, Caroline Aubé, and Jean Talbot. 2021. The Influence of Video Format on Engagement and Performance in Online Learning. Brain Sciences 11 (2021). https://doi.org/10.3390/brainsci11020128

[45] Joseph J LaViola Jr, Ernst Kruijff, Ryan P McMahan, Doug Bowman, and Ivan P Poupyrev. 2017. 3D user interfaces: theory and practice. Addison-Wesley Professional.

[46] Benjamin Lee, Dave Brown, Bongshin Lee, Christophe Hurter, Steven Drucker, and Tim Dwyer. 2021. Data Visceralization: Enabling Deeper Understanding of Data Using Virtual Reality. IEEE Transactions on Visualization and Computer Graphics 27, 2 (2021), 1095–1105. https://doi.org/10.1109/TVCG.2020.3030435

[47] Geonsun Lee, HyeongYeop Kang, JongMin Lee, and JungHyun Han. 2020. A User Study on View-sharing Techniques for One-to-Many Mixed Reality Collaborations. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). 343–352. https://doi.org/10.1109/VR46266.2020.00054

[48] Jiwon Lee, Mingyu Kim, and Jinmo Kim. 2020. RoleVR: Multi-Experience in Immersive Virtual Reality between Co-Located HMD and Non-HMD Users. Multimedia Tools Appl. 79, 1–2 (jan 2020), 979–1005. https://doi.org/10.1007/s11042-019-08220-w

[49] Jian Liao, Adnan Karim, Shivesh Singh Jadon, Rubaiat Habib Kazi, and Ryo Suzuki. 2022. RealityTalk: Real-Time Speech-Driven Augmented Presentation for AR Live Storytelling. In Proceedings of the 35th Annual ACM Symposium on User Interface Software and Technology (Bend, OR, USA) (UIST '22). Association for Computing Machinery, New York, NY, USA, Article 17, 12 pages. https://doi.org/10.1145/3526113.3545702

[50] Gustavo Lopez and Luis A. Guerrero. 2017. Awareness Supporting Technologies used in Collaborative Systems: A Systematic Literature Review. In Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing (Portland, Oregon, USA) (CSCW '17). Association for Computing Machinery, New York, NY, USA, 808–820. https://doi.org/10.1145/2998181.2998281

[51] Christian Mai, Lukas Rambold, and Mohamed Khamis. 2017. TransparentHMD: revealing the HMD user's face to bystanders. In Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia (Stuttgart, Germany) (MUM '17). Association for Computing Machinery, New York, NY, USA, 515–520. https://doi.org/10.1145/3152832.3157813

[52] Nathan Matsuda, Brian Wheelwright, Joel Hegland, and Douglas Lanman. 2021. VR social copresence with light field displays. ACM Trans. Graph. 40, 6, Article 244 (dec 2021), 13 pages. https://doi.org/10.1145/3478513.3480481

[53] Gerard McAtamney and Caroline Parker. 2006. An examination of the effects of a wearable display on informal face-to-face communication. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (Montréal, Québec, Canada) (CHI '06). Association for Computing Machinery, New York, NY, USA, 45–54. https://doi.org/10.1145/1124772.1124780

[54] 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 (Seoul, Republic of Korea) (CHI '15). Association for Computing Machinery, New York, NY, USA, 2143–2152. https://doi.org/10.1145/2702123.2702382

[55] 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, Hari Das (Ed.), Vol. 2351. International Society for Optics and Photonics, SPIE, 282 – 292. https://doi.org/10.1117/12.197321

[56] Teresa Monahan, Gavin McArdle, and Michela Bertolotto. 2008. Virtual reality for collaborative e-learning. Computers & Education 50, 4 (2008), 1339–1353.

https://doi.org/10.1016/j.compedu.2006.12.008

[57] Catarina Moreira, Francisco P. M. Simões, Mark J. W. Lee, Ezequiel R. Zorzal, Robert W. Lindeman, João Madeiras Pereira, Kyle Johnsen, and Joaquim Jorge. 2023. Toward VR in VR: Assessing Engagement and Social Interaction in a Virtual Conference. IEEE Access 11 (2023), 1906–1922. https://doi.org/10.1109/ACCESS.2022.3233312

[58] Prasanth Murali, Javier Hernandez, Daniel McDuff, Kael Rowan, Jina Suh, and Mary Czerwinski. 2021. AffectiveSpotlight: Facilitating the Communication of Affective Responses from Audience Members during Online Presentations. 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 247, 13 pages. https://doi.org/10.1145/3411764.3445235

[59] Nels Numan and Anthony Steed. 2022. Exploring User Behaviour in Asymmetric Collaborative Mixed Reality. In Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology (Tsukuba, Japan) (VRST '22). Association for Computing Machinery, New York, NY, USA, Article 6, 11 pages. https://doi.org/10.1145/3562939.3565630

[60] Ken Perlin, Zhenyi He, and Karl Rosenberg. 2018. Chalktalk : A Visualization and Communication Language – As a Tool in the Domain of Computer Science Education. https://doi.org/10.48550/arXiv.1809.07166

[61] Danakorn Nincarean Eh Phon, Mohamad Bilal Ali, and Noor Dayana Abd Halim. 2014. Collaborative Augmented Reality in Education: A Review. In 2014 International Conference on Teaching and Learning in Computing and Engineering. 78–83. https://doi.org/10.1109/LaTiCE.2014.23

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

[63] Thammathip Piumsomboon, Gun A. Lee, Andrew Irlitti, Barrett Ens, Bruce H. Thomas, and Mark Billinghurst. 2019. On the Shoulder of the Giant: A Multi-Scale Mixed Reality Collaboration with 360 Video Sharing and Tangible Interaction. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (Glasgow, Scotland Uk) (CHI '19). Association for Computing Machinery, New York, NY, USA, 1–17. https://doi.org/10.1145/3290605.3300458

[64] 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 (April 2021), 21 pages. https://doi.org/10.1145/3449243

[65] Troels A. Rasmussen and Weidong Huang. 2019. SceneCam: Using AR to improve Multi-Camera Remote Collaboration. In SIGGRAPH Asia 2019 XR (Brisbane, QLD, Australia) (SA '19). Association for Computing Machinery, New York, NY, USA, 36–37. https://doi.org/10.1145/3355355.3361892

[66] Nico Reski, Aris Alissandrakis, Jukka Tyrkkö, and Andreas Kerren. 2020. "Oh, that's where you are!" – Towards a Hybrid Asymmetric Collaborative Immersive Analytics System. In Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society (Tallinn, Estonia) (NordiCHI '20). Association for Computing Machinery, New York, NY, USA, Article 5, 12 pages. https://doi.org/10.1145/3419249.3420102

[67] Mauro Del Rio, Vittorio Meloni, Francesca Frexia, Francesco Cabras, Roberto Tumbarello, Sabrina Montis, Andrea Marini, and Gianluigi Zanetti. 2018. Augmented Reality for Supporting Real Time Telementoring: an Exploratory Study Applied to Ultrasonography. In Proceedings of the 2nd International Conference on Medical and Health Informatics (Tsukuba, Japan) (ICMHI '18). Association for Computing Machinery, New York, NY, USA, 218–222. https://doi.org/10.1145/3239438.3239444

[68] Hans Rosling. 2010. 200 Countries, 200 Years, 4 Minutes. BBC Four. https://www.youtube.com/watch?v=jbkSRLYSojo

[69] Louis F Ruado and Leah Amor S Cortez. 2024. Enhancing Student Engagement and Achievement in Biology through Interactive Slide Presentations. American Journal of Education and Technology 3, 1 (Mar. 2024), 51–59. https://doi.org/10.54536/ajet.v3i1.2520

[70] Grace V Ryan, Shauna Callaghan, Anthony Rafferty, Mary F Higgins, Eleni Mangina, and Fionnuala McAuliffe. 2022. Learning Outcomes of Immersive Technologies in Health Care Student Education: Systematic Review of the Literature. J Med Internet Res 24, 2 (1 Feb 2022), e30082. https://doi.org/10.2196/30082

[71] David Saffo, Andrea Batch, Cody Dunne, and Niklas Elmqvist. 2023. Through Their Eyes and In Their Shoes: Providing Group Awareness During Collaboration Across Virtual Reality and Desktop Platforms. In Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23). Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3544548.3581093

[72] David Saffo, Sara Di Bartolomeo, Tarik Crnovrsanin, Laura South, Justin Raynor, Caglar Yildirim, and Cody Dunne. 2024. Unraveling the Design Space of Immersive Analytics: A Systematic Review. IEEE Transactions on Visualization and Computer Graphics 30, 1 (2024), 495–506. https://doi.org/10.1109/TVCG.2023.3327368

[73] Nazmus Saquib, Rubaiat Habib Kazi, Li-Yi Wei, and Wilmot Li. 2019. Interactive Body-Driven Graphics for Augmented Video Performance. 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, 1–12. https://doi.org/10.1145/3290605.3300852

[74] Prasanth Sasikumar, Lei Gao, Huidong Bai, and Mark Billinghurst. 2019. Wearable RemoteFusion: A Mixed Reality Remote Collaboration System with Local Eye Gaze and Remote Hand Gesture Sharing. In *2019 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*. 393–394. https://doi.org/10.1109/ISMAR-Adjunct.2019.000-3

[75] Joachim Scholz and Andrew N. Smith. 2016. Augmented reality: Designing immersive experiences that maximize consumer engagement. *Business Horizons* 59, 2 (2016), 149–161. https://doi.org/10.1016/j.bushor.2015.10.003

[76] Jan-Henrik Schröder, Daniel Schacht, Niklas Peper, Anita Marie Hamurculu, and Hans-Christian Jetter. 2023. Collaborating Across Realities: Analytical Lenses for Understanding Dyadic Collaboration in Transitional Interfaces. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 97, 16 pages. https://doi.org/10.1145/3544548.3580879

[77] Kyvete Shatri and Lediana Shala. 2022. Evaluating the Effect of Interactive Digital Presentations on Students' Performance during Technology Class. *Education Research International* 2022, 1 (2022), 3337313. https://doi.org/10.1155/2022/3337313

[78] Donghee Shin. 2019. How does immersion work in augmented reality games? A user-centric view of immersion and engagement. *Information, Communication & Society* 22, 9 (2019), 1212–1229. https://doi.org/10.1080/1369118X.2017.1411519

[79] Vivek K. Singh, Ankur Mani, and Alex Pentland. 2014. Social Persuasion in Online and Physical Networks. *Proc. IEEE* 102, 12 (2014), 1903–1910. https://doi.org/10.1109/JPROC.2014.2363986

[80] Richard Skarbez, Missie Smith, and Mary C. Whitton. 2021. Revisiting Milgram and Kishino's Reality-Virtuality Continuum. *Frontiers in Virtual Reality* 2 (2021). https://doi.org/10.3389/frvir.2021.647997

[81] Mel Slater, R. Lotto, Maria Arnold, and Maria Sanchez-Vives. 2009. How we experience immersive virtual environments: The concept of presence and its measurement. *Anuario de Psicologia* 40 (09 2009).

[82] Jim Smiley, Benjamin Lee, Siddhant Tandon, Maxime Cordeil, Lonni Besançon, Jarrod Knibbe, Bernhard Jenny, and Tim Dwyer. 2021. The MADE-Axis: A Modular Actuated Device to Embody the Axis of a Data Dimension. *Proceedings of the ACM on Human-Computer Interaction* 5, ISS (Nov. 2021), 1–23. https://doi.org/10.1145/3488546

[83] Mark St. John, Michael B. Cowen, Harvey S. Smallman, and Heather M. Oonk. 2001. The Use of 2D and 3D Displays for Shape-Understanding versus Relative-Position Tasks. *Human Factors* 43, 1 (2001), 79–98. https://doi.org/10.1518/001872001775992534

[84] Dietrich Stoevesandt, Andreas Weber, Andreas Wienke, Steffi Bethge, Viktoria Heinze, Simone Kowoll, and Axel Schlitt. 2020. Interactive patient education via an audience response system in cardiac rehabilitation. *SAGE Open Medicine* 8 (2020), 10 pages. https://doi.org/10.1177/2050312120942118

[85] Zhiyao Sun, Tian Lv, Sheng Ye, Matthieu Lin, Jenny Sheng, Yu-Hui Wen, Minjing Yu, and Yong-Jin Liu. 2024. DiffPoseTalk: Speech-Driven Stylistic 3D Facial Animation and Head Pose Generation via Diffusion Models. *ACM Transactions on Graphics (TOG)* 43, 4, Article 46 (2024), 9 pages. https://doi.org/10.1145/3658221

[86] Kartik Teotia, Mallikarjun B R, Xingang Pan, Hyeongwoo Kim, Pablo Garrido, Mohamed Elgharib, and Christian Theobalt. 2024. HQ3DAvatar: High-quality Implicit 3D Head Avatar. *ACM Trans. Graph.* 43, 3, Article 27 (Apr 2024), 24 pages. https://doi.org/10.1145/3649889

[87] 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

[88] M. Tory, A.E. Kirkpatrick, M.S. Atkins, and T. Moller. 2006. Visualization task performance with 2D, 3D, and combination displays. *IEEE Transactions on Visualization and Computer Graphics* 12, 1 (2006), 2–13. https://doi.org/10.1109/TVCG.2006.17

[89] Thomas Tøth. 2014. Trust in Client-Vendor Relations: An Empirical Study of Collaboration across National and Organizational Boundaries. In *Proceedings of the 5th ACM International Conference on Collaboration across Boundaries: Culture, Distance & Technology* (Kyoto, Japan) *(CABS '14)*. Association for Computing Machinery, New York, NY, USA, 5–14. https://doi.org/10.1145/2631488.2631491

[90] Marco Viola. 2022. Seeing through the shades of situated affectivity. Sunglasses as a socio-affective artifact. *Philosophical Psychology* (2022), 1–25. https://doi.org/10.1080/09515089.2022.2118574

[91] Amy Voida, Stephen Voida, Saul Greenberg, and Helen Ai He. 2008. Asymmetry in media spaces. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (San Diego, CA, USA) *(CSCW '08)*. Association for Computing Machinery, New York, NY, USA, 313–322. https://doi.org/10.1145/1460563.1460615

[92] Jan B. Vornhagen, April Tyack, and Elisa D. Mekler. 2020. Statistical Significance Testing at CHI PLAY: Challenges and Opportunities for More Transparency. In *Proceedings of the Annual Symposium on Computer-Human Interaction in Play* (Virtual Event, Canada) *(CHI PLAY '20)*. Association for Computing Machinery, New York, NY, USA, 4–18. https://doi.org/10.1145/3410404.3414229

[93] Xue Wang, Lin Lin, Meiqi Han, and J. Michael Spector. 2020. Impacts of cues on learning: Using eye-tracking technologies to examine the functions and designs of added cues in short instructional videos. *Computers in Human Behavior* 107 (2020), 106279. https://doi.org/10.1016/j.chb.2020.106279

[94] Jane Webster and Hayes Ho. 1997. Audience engagement in multimedia presentations. *SIGMIS Database* 28, 2 (Apr 1997), 63–77. https://doi.org/10.1145/264701.264706

[95] Gareth White. 2012. On Immersive Theatre. *Theatre Research International* 37, 3 (2012), 221–235. https://doi.org/10.1017/S0307883312000880

[96] Wesley Willett, Yvonne Jansen, and Pierre Dragicevic. 2017. Embedded Data Representations. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2017), 461–470. https://doi.org/10.1109/TVCG.2016.2598608

[97] Jeanne M. Wilson, Susan G. Straus, and Bill McEvily. 2006. All in due time: The development of trust in computer-mediated and face-to-face teams. *Organizational Behavior and Human Decision Processes* 99, 1 (2006), 16–33. https://doi.org/10.1016/j.obhdp.2005.08.001

[98] Yalong Yang, Bernhard Jenny, Tim Dwyer, Kim Marriott, Haohui Chen, and Maxime Cordeil. 2018. Maps and Globes in Virtual Reality. *Computer Graphics Forum* 37, 3 (2018), 427–438. https://doi.org/10.1111/cgf.13431

[99] Lei Zhang, Daekun Kim, Youjean Cho, Ava Robinson, Yu Jiang Tham, Rajan Vaish, and Andrés Monroy-Hernández. 2024. Jigsaw: Authoring Immersive Storytelling Experiences with Augmented Reality and Internet of Things. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '24)*. Association for Computing Machinery, New York, NY, USA, Article 87, 14 pages. https://doi.org/10.1145/3613904.3642744

[100] Qingcheng Zhao, Pengyu Long, Qixuan Zhang, Dafei Qin, Han Liang, Longwen Zhang, Yingliang Zhang, Jingyi Yu, and Lan Xu. 2024. Media2Face: Co-speech Facial Animation Generation With Multi-Modality Guidance. In *ACM SIGGRAPH 2024 Conference Papers* (Denver, CO, USA) *(SIGGRAPH '24)*. Association for Computing Machinery, New York, NY, USA, Article 18, 13 pages. https://doi.org/10.1145/3641519.3657413

A Supplementary Materials

This supplementary material provides tables showing the specific questionnaire questions and more fine-grained results of participants' questionnaire responses and the behaviors we observed.

A.1 Questionnaires

Category	Element	Question
What	Action	I always knew what my partner was doing.
	Intention	I always understood the goal of my partner's actions.
	Artifact - Partner	When they referred to something, I always knew what object my partner was referring to.
	Artifact - Self	My partner could always understand what I was referring to.
Where	Location	I always knew where my partner was located.
	Gaze	I always knew where my partner was looking.
	View	I always knew what my partner could see.
	Reach	I always knew where my partner could reach.

Table 1: Group awareness questionnaire.

Category	Question (presenter adaptation)
Attention Focus	This presentation medium keeps me (my audience) totally absorbed in the presentation. This presentation medium holds my (audience's) attention.
Curiosity	This presentation medium excites my (audience's) curiosity. This presentation medium arouses my (audience's) imagination.
Intrinsic Interest	This presentation medium is fun (for my audience). This presentation medium is intrinsically interesting (for my audience).
Overall	This presentation medium is engaging (for my audience).
Content	I was engaged (was able to engage my audience) with the presentation's virtual content.
Interpersonal	I was engaged with the presenter themselves. (I was able to engage my audience interpersonally/socially.)

Table 2: Presentation engagement questionnaire.

A.2 Questionnaire Responses

In these tables, IQR indicates the interquartile range of the data. Q1 and Q3 represent the first quartile (25%) and the third quartile (75%) of the data, respectively.

Subscale	Modality	Median	IQR	Q1	Q3
Mental Demand	Symmetric	3.50	2.50	2.25	4.75
	Asymmetric	3.00	2.25	1.50	3.75
Physical Demand	Symmetric	2.50	2.50	1.25	3.75
	Asymmetric	2.50	1.75	1.25	3.00
Performance	Symmetric	5.50	0.75	5.25	6.00
	Asymmetric	5.50	1.75	5.00	6.75

Table 3: NASA-TLX results.

Modality	Participant Type	Median	IQR	Q1	Q3
Symmetric	Presenter	5.75	2.03	4.22	6.25
	Audience	6.12	0.34	6.00	6.34
Asymmetric	Presenter	5.31	0.72	5.00	5.72
	Audience	4.62	0.38	4.53	4.91

Table 4: Group awareness questionnaire responses.

Modality	Participant Type	Median	IQR	Q1	Q3
Symmetric	Presenter	6.39	1.11	5.75	6.86
	Audience	6.50	0.53	6.36	6.89
Asymmetric	Presenter	6.06	1.25	5.47	6.72
	Audience	5.94	0.77	5.56	6.33

Table 5: Engagement questionnaire responses.

A.3 Observational Data

In these tables, IQR indicates the interquartile range of the data. Q1 and Q3 represent the first quartile (25%) and the third quartile (75%) of the data, respectively.

Modality	Participant Type	Median	IQR	Q1	Q3
Symmetric	Presenter	0.50	2.50	0.00	2.50
	Audience	2.50	1.75	1.25	3.00
Asymmetric	Presenter	1.50	4.50	0.00	4.50
	Audience	4.50	13.50	0.00	13.50

Table 6: Counts of social expressions observed.

Modality	Participant Type	Median (%)	IQR (%)	Q1 (%)	Q3
Symmetric	Presenter	5.65	8.30	1.65	9.95
	Audience	5.45	8.20	3.80	12.00
Asymmetric	Presenter	6.65	10.05	1.15	11.20
	Audience	2.80	4.33	1.35	5.68

Table 7: Percentage of presentation participants looked at each other.

Modality	Median	IQR	Q1	Q3
Symmetric	2.00	3.50	1.00	4.50
Asymmetric	2.50	3.25	0.50	3.75

Table 8: Counts of occurrences at which participants made eye contact with each other.

Modality	Median	IQR	Q1	Q3
Symmetric	0.38	2.52	0.08	2.60
Asymmetric	1.00	2.47	0.18	2.65

Table 9: Percentage of the presentation that participants made eye contact with each other.

Modality	Label	Median	IQR	Q1	Q3
Symmetric	Presenter	11.50	13.75	7.25	21.00
	Audience	5.50	1.00	5.00	6.00
Asymmetric	Presenter	8.50	6.80	4.00	10.80
	Audience	7.00	5.25	2.50	7.75

Table 10: Counts of total references to presentation content participants made.

Reference	Modality	Median	IQR	Q1	Q3
Pointing	Symmetric	6.00	7.45	3.75	11.20
	Asymmetric	3.50	6.25	1.50	7.75
Deictic Speech	Symmetric	5.50	8.25	1.00	9.25
	Asymmetric	1.50	1.00	1.00	2.00
Relative Language	Symmetric	0.00	0.75	0.00	0.75
	Asymmetric	1.00	1.50	0.25	1.75
Verbal Property References	Symmetric	0.00	0.75	0.00	0.75
	Asymmetric	1.00	0.50	0.25	0.75
Verbal Reference Chaining	Symmetric	0.00	0.00	0.00	0.00
	Asymmetric	0.00	0.00	0.00	0.00

Table 11: Counts of reference types by presenter modality.

Modality	Behavior	Participant Type	Median	IQR	Q1	Q3
Asymmetric	Pointing	Audience	4.50	4.00	1.75	5.75
		Presenter	3.50	6.25	1.50	7.75
	Deictic speech	Audience	1.50	3.00	0.00	3.00
		Presenter	1.50	1.00	1.00	2.00
	Relative language	Audience	0.00	0.00	0.00	0.00
		Presenter	1.00	1.50	0.25	1.75
	Verbal property reference	Audience	0.00	0.00	0.00	0.00
		Presenter	0.00	3.00	0.00	3.00
	Referential chaining	Audience	0.00	0.00	0.00	0.00
		Presenter	0.00	0.00	0.00	0.00
Symmetric	Pointing	Audience	3.00	3.50	1.25	4.75
		Presenter	6.00	7.50	3.75	11.25
	Deictic speech	Audience	1.00	1.50	0.25	1.75
		Presenter	5.50	8.25	1.00	9.25
	Relative language	Audience	0.00	0.00	0.00	0.00
		Presenter	0.00	0.75	0.00	0.75
	Verbal property reference	Audience	0.00	0.75	0.00	0.75
		Presenter	0.00	0.75	0.00	0.75
	Referential chaining	Audience	0.00	0.00	0.00	0.00
		Presenter	0.00	0.00	0.00	0.00

Table 12: Counts of behavior types observed during presentations.

Disclaimer