

Investigating the Influence of Playback Interactivity during Guided Tours for Asynchronous Collaboration in Virtual Reality

Alexander Giovannelli^{1*}
Vuthea Chheang²

Leonardo Pavanatto¹
Brian Giera²

Shakiba Davari¹
Timo Bremer²

Haichao Miao²
Doug A. Bowman¹

¹ Center for Human-Computer Interaction, Virginia Tech, USA

² Lawrence Livermore National Laboratory, USA

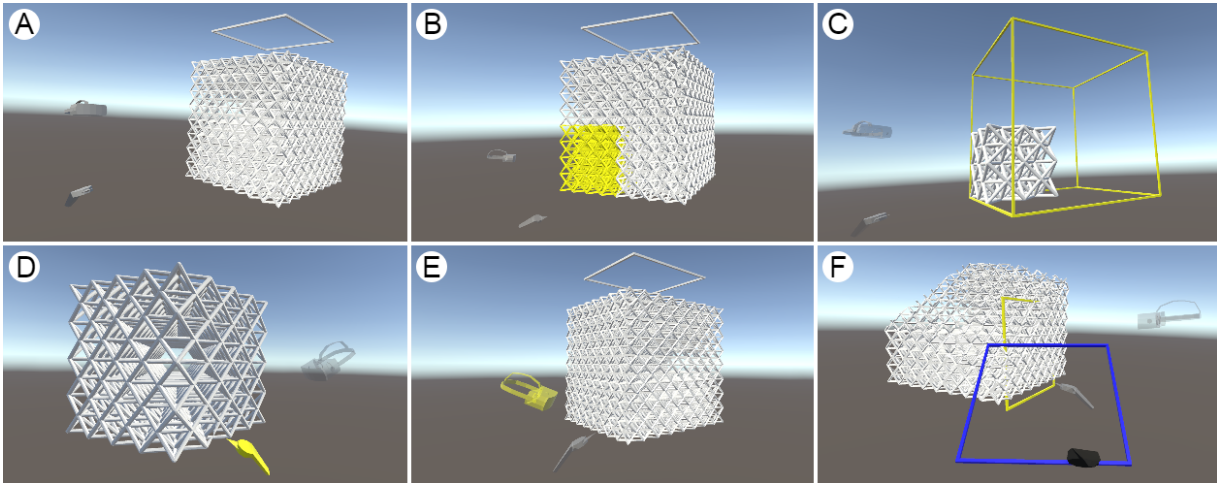


Figure 1: View of the inspection environment from the participant’s field of view with corresponding yellow pause interaction indicators. The guide is the gray avatar, whose opacity is reduced during a pause requiring a given playback interaction. A: Standard playback. B: *Navigation Down* interaction. C: *Navigation Up* interaction. D: *Controller Point* interaction. E: *WYSIWIS* interaction. F: *Plane Dock* interaction.

ABSTRACT

Collaborative virtual environments allow workers to contribute to team projects across space and time. While much research has closely examined the problem of working in different spaces at the same time, few have investigated the best practices for collaborating in those spaces at different times aside from textual and auditory annotations. We designed a system that allows experts to record a tour inside a virtual inspection space, preserving knowledge and providing later observers with insights through a 3D playback of the expert’s inspection. We also created several interactions to ensure that observers are tracking the tour and remaining engaged. We conducted a user study to evaluate the influence of these interactions on an observing user’s information recall and user experience. Findings indicate that independent viewpoint control during a tour enhances the user experience compared to fully passive playback and that additional interactivity can improve auditory and spatial recall of key information conveyed during the tour.

Index Terms: Virtual reality, computer-supported cooperative work, collaborative virtual environments, asynchronous collaboration, guided tours.

*e-mail: agiovannelli@vt.edu

1 INTRODUCTION

As Virtual Reality (VR) head-worn displays (HWDs) become more commercially viable in terms of cost and quality, designers and developers must consider the technology’s most impactful application spaces. An area of interest over the years has been collaboration within immersive environments, where a plethora of new cooperative capabilities can be realized [16]. Coined as Collaborative Virtual Environments (CVEs), these digital landscapes allow multiple users to communicate with one another and interact with digital content through various interaction methods [8, 47, 53]. As a result of the geographical distribution of work becoming more common [54] in recent years, telepresence tools such as Zoom¹ and WebEx² have been used to facilitate remote collaboration. Although these desktop-based collaboration applications help facilitate basic screen sharing and whiteboarding, they forfeit the benefit of embodied interaction, representation of information, and collaborators that CVEs provide for cooperative work [37, 17].

While a large amount of literature has focused on addressing the challenges of CVEs in the realm of synchronous collaboration [11], using CVEs for asynchronous collaboration has been identified as an area of limited exploration [24, 43]. CVEs have the potential to preserve increments of work [63], allow the annotation of knowledge and reasoning from individual contributions [34], draw collaborator attention and awareness to these annotations [51], and enable ad-hoc communicative exchange and relaying of pertinent information to drive continuous work across all hours [35]. These innate

¹<https://zoom.us/>

²<https://www.webex.com/>

features of CVEs make them rich for further research investigation in their applicability for asynchronous collaboration.

We propose extending asynchronous collaboration research in VR via guided tours. Guided tours create an observer-guide relationship within immersive environments to allow narrative relaying of information across both space and time [25]. Within a CVE, this would empower the preservation of expert knowledge regarding changes inside the environment made during a working session or the creation of overview information for novice observing users to follow later. We motivate our work in the additive manufacturing (AM) specialty, falling within previously identified areas of potential asynchronous collaboration in engineering [37]. During AM, inspection processes occur regularly, requiring experts across design, engineering, and manufacturing areas to provide feedback to ensure the quality of the produced part [5]. However, whether from geographical distribution or scheduling constraints due to other responsibilities, these experts may not be able to inspect a given part jointly. By using an inspection CVE, these experts can guide later observers through their inspection findings, preserving their analyses for continued collaboration asynchronously.

While in some respects viewing guided tours in a CVE is like a synchronous collaboration, the observer cannot directly communicate with the tour guide, nor can they influence the content of the tour. This is analogous to the difference between attending an in-person lecture versus viewing a prerecorded lecture video. Thus, we expect there might be a higher likelihood for observers to lose engagement with the tour, or to fail to completely follow the tour's informational content. We hypothesize that embedding opportunities for the observer to interact with the playback or the environment during the tour will mitigate these effects and will lead to better information retention and user experience.

This paper details our design of a guided tour system with playback interactivity and a subsequent user study investigating its use for asynchronous collaboration in the context of AM. We detail our findings regarding the use of interaction as a means to view and control the playback of the guided tour and its influence on the user's experience and retention of information covered in the tour's narrative. We synthesize the findings with participant feedback and open-ended discussion, supporting the use of interactivity during playback when conducting guided tours instead of passive viewership. We also document the perceived benefits of interaction for playback control. Contributions of this work include (1) the design of interactive methods to maintain observer engagement during guided tours in a CVE; (2) quantitative measurement of four varying increments of interactivity on user recall and user experience using a guided tour playback system; and (3) qualitative feedback regarding the design of interactive behaviors in a guided tour playback system.

2 RELATED WORK

2.1 Guided Tours

Using "The River Analogy," Gaylean proposed author-controlled navigation of virtual environments (VEs) to empower the structuring of experiences for informative presentation [14]. The analogy itself consisted of navigation akin to riding a boat on a river, where the users on the boat have minor control of direction, but ultimately the author controls the current by which the users move downstream and are shown the surrounding landscape. While this did present a novel means of presenting information, the extent of control for users was reduced to allow guiding with no interruption.

Pausch et al. highlighted that such an approach would detract from the overall experience, where users would be unimpressed by the technology being used for its own sake; they care more about what there is that they can do in the virtual world [40]. Through developing a VR system and evaluating it at a theme park, they found that most users preferred complete control of navigation, navigating

to their own pointed places of interest in the experience. This same user-controlled approach was applied in a PC-based experience for physicians to tour a patient's colon, allowing them to perform non-invasive colonoscopy procedures [22]. The usage of guided navigation was found to empower computer-illiterate surgeons to successfully inspect areas of interest inside the colon by instead using mouse peripherals to view locations of interest.

Best further extended the concept of guided tours in the context of museums, proposing digital guides to deliver personalized tours for users [3]. These digital guides would act as a pathfinder, leading users around a unique site, and a mentor, providing key information to users about that site. However, Best's insights were gleaned from her own observations of recorded, real-world museum tours, where Tsiropoulou et al. proposed a human-in-the-loop approach [55]. They highlight the importance of capturing the reported mental, physical, and spatial impacts of visitors' overall quality of experience obtained from visiting individual exhibits.

From this existing literature, work regarding guided tours has primarily focused on controlling the presentation of data or knowledge by managing the extent of an individual's ability to navigate and interact inside virtual and physical space. While the potential use of digital guides to steer observers to points of interest and provide key information is addressed, no resulting evaluation or use is investigated. We explore this digital guide concept in our work as a means to not only control the flow of information shared but also to measure the impact on recall and experience.

2.2 Annotations

Early usage of VEs focused on their capability to not only serve as simulator systems but to empower the marking of a VE to preserve information for future review by a peer or collaborator using annotations. Verlinden et al. described such a system, noting the possibilities of using annotations to link temporal features of a VE [56]. They noted the importance of documenting findings associated with various data objects such that "notes are spatially indexed" and that, within the simulation, those annotations should be "natural" and create new ways to communicate with others. A similar system was introduced by Harmon et al., adding the aforementioned annotation association capabilities by supporting "object annotation" and "view annotation" [20].

Frécon & Nöu leveraged these annotation methods explicitly for collaboration by creating a VE modeled after a meeting room and providing a set of virtual tools allowing users to store text artifacts (e.g., notebooks and reports) and further annotate those artifacts [13]. Since these early works, annotations have been investigated thoroughly, identifying characteristics such as placement methods [4, 60, 42] and media types of text [32, 50], speech [56, 44], images [49], videos [19], sketches [18, 9], and even world-views [2].

Although these researched systems and types of annotations provide features that theoretically support collaboration, they do not explicitly investigate the implications of use for asynchronous collaboration. In our work, we investigate using guiding avatars as a means to annotate information in the "natural" way.

2.3 Asynchronous Collaboration

Asynchronous collaboration's flexibility to allow contribution of work regardless of time is its greatest strength compared to its synchronous counterpart [23]. However, while research has largely solved the authoring of annotations for asynchronous collaboration, the complexity of reviewing these annotations at a later time and the resulting understanding from a review has been seldom investigated [24]. Although identified as a topic of limited research interest in VR [43], asynchronous collaboration has seen many new contributions in recent years.

Wang et al. developed a communication tool in VR for 3D asynchronous collaborative design, comparing it with desktop-audio and desktop-based CVEs [57]. The system recorded and replayed avatars as they placed and moved furniture within a room, with participants preferring the use of VR in dimensions of communication clarity, perception of partners, performance satisfaction, and outcome satisfaction. Chow et al. similarly contributed to the area of interior design, further analyzing the social behavior in asynchronous collaboration [7]. They proposed design recommendations from their evaluation: providing rich navigation cues, activity highlighting, and animation changes at transitions in the CVE.

Outside the interior design and decorating use case, Marques et al. examined the use of notification sharing via spatial annotations in maintenance scenarios with augmented reality (AR) [34, 35]. Their system allowed onsite technicians to use a handheld device to annotate problems for remote experts to instruct them later on repair methods. García-Pereira created a similar system, focusing on the quality of handheld devices used to create spatial annotations and to preserve position in search and find tasks for later users [15]. Lee et al. also explored this concept of using AR spatial annotations to provide instructional videos on everyday tasks [29]. Several works have supplemented these works by analyzing methods to assist in the authorship of spatial annotations, with Wang et al. exploring the use of freehand behaviors for context-aware creation and responsiveness of AR virtual contents [58, 59].

Expanding to asynchronous collaboration across the reality-virtuality continuum [38], Fender et al. created and analyzed a system that preserved co-located or remote interruptions from an observer of an occupied VR user [12]. A similar system was produced by Cho et al. for tracking and visualizing events in a mixed reality user's physical space to provide enhancements to memory awareness of these background occurrences [6].

In the manufacturing space, Mayer et al. developed and evaluated a system that had a presenting avatar perform an assembly process [36]. Participants could follow this collaborative avatar as opposed to learning from a virtual manual to complete their own assembly. They further noted the potential of using asynchronous collaboration systems in the digitalization of artifacts to empower remote work in business, education, and engineering sectors [37].

While these works have examined asynchronous collaboration using VR technologies, they have focused on social outcomes, question-and-answer behaviors, and the performing of instructions with explicit playback controls. Our work builds on this by evaluating the use of interactivity as a means to maintain engagement and understanding during an asynchronous inspection process.

3 GUIDED TOUR PLAYBACK INTERACTION DESIGN

We designed a system for a guided tour where a mentoring guide explained the inspection process of a 3D-printed object. The object within the VE was generated using open-source computed tomography data from AM octet lattice structures³ and the narrative spoken by the guide from the literature regarding these structures [28]. During the tour, the guide's voice was heard explaining the inspection verbally, and the guide's avatar (i.e., head and hands as shown in Fig. 1A) was seen moving around the object, looking or pointing at various parts of the object, cutting into the object with a cutting plane, and navigating through different levels of scale of the object.

Our goal was to foster the observing user's engagement in the guided tour during the inspection process. To achieve this goal, we designed interactions to make the user mimic the guide's actions during the tour. To integrate these interactions into the tour, the playback would pause whenever the guide performed any of the

following: navigation across scales, usage of tools, and view of regions or cells of interest on the 3D-printed object. The playback would resume once the user performed the corresponding interaction based on the guide's actions. The guide is responsible for specifying the interactions they want the observing user to mimic by using explicit controller input in a separate recording mode available to designated guide users.

For these playback interactions, visual and auditory feedback indicated to the user that an interaction was required. The audio cue consisted of a unique sound per interaction and was played at the user's position. The visual indicator included an object in the environment spawning with a yellow material, highlighting the required interaction for continued playback. Additionally, the guide's avatar turned from gray to transparent gray while the playback was paused, returning to gray once the required interaction was performed and playback resumed. The following details the mimicking interactions in generalized groups.

Multiscale Navigation During the inspection, the guide traverses across scales of the 3D-printed object, effectively zooming in and out of specified sections. This provides a closer, detailed view of a smaller group of cells comprising the print versus a wider, less detailed view of many cells. We divide multiscale navigation capability across two interactions: *Navigate Down* and *Navigate Up*.

Navigate Down: When the guide navigates down, their avatar shrinks in scale by a multiplier of 0.5. The octant they selected for the navigation down then becomes yellow, as shown in Fig. 1B. The user must then perform the same narrowing navigation by aiming at the highlighted octant and pressing the controller's trigger to continue the tour playback. This reduces the player's scale by the 0.5 multiplier and removes the cells outside of the selected octant to disappear from the environment. The number of unit cells displayed when navigating down in scale is 512 to 64 and 64 to 8.

Navigate Up: When the guide navigates up, their avatar grows in scale by a multiplier of 2. The position in the print that they have navigated up toward is represented by a yellow box, as shown in Fig. 1C. The user must then perform the same upward navigation by pressing the primary button on the controller. This grows the player's scale by the multiplier of 2, populates the yellow grid with the cells comprising it, and resumes the tour playback. The number of unit cells displayed when navigating up in scale is 8 to 64 and 64 to 512.

Perspective Sharing During the inspection, the guide points toward and looks at regions, features, or specific cells comprising the 3D-printed object, calling for the user to pay attention to them. This causes the user to identify these features or specific cells by visually directing their attention toward them before a verbal explanation is given describing their significance. We divide the visuospatial relation of described features or cells across three interactions: *Controller Point*, *WYSIWIS*, and *Plane Dock*.

Controller Point: Whenever the guide pointed to a region or cell of interest that required the user's attention, the guide's controller would stop before that region or cell and turn yellow, as shown in Fig. 1D. The user must then observe the point of interest specified by the guide's controller. They then aimed at the guide's controller, pressing their own controller's trigger to acknowledge the point of interest and resume tour playback.

WYSIWIS: When the guide wanted the user to observe the 3D-printed object from a specific point of view, the guide's head would stop at that viewpoint and turn a transparent yellow, as shown in Fig. 1E. The user then needed to position their head within the guide avatar's head to see from their perspective, aligning with its approximate rotation for 1.5 seconds. After completing this interaction, known as "What You See Is What I See", the tour playback resumed [48].

³<https://data-science.llnl.gov/open-data-initiative>

Test A	Test B	Question Type
What is the potential cause of bent and broken strut defects?	How many unit cells exist in the second level of the model?	Verbal
Place the red sphere at the location of the second defect discussed in the tour.	Place the red sphere at the location of the Bent strut defect.	Spatial
What is the second defect referred to as?	Which defect is the most severe?	Verbal
Place the red sphere at the location of the Broken strut defect.	Place the red sphere at the location of the Thin strut defect.	Spatial
How does a 3D printer create a thin strut defect?	What are the individual unit cells referred to as?	Verbal

Table 1: Test questions in Test A and Test B.

Plane Dock: A cutting plane in the environment allowed the guide to cut through the 3D-printed object, revealing its inner structure, removing occluding cells, and providing a desired view of the remaining cells. To have the user understand the positioning of the plane and the resulting desired view of the cells, the guide spawned a yellow dock at the position and rotation of a required cut, as shown in Fig. 1F. The user then needed to grab the cutting plane by hovering over it with their controller and pressing the trigger on the controller, then placing it within the approximate position and rotation of the yellow dock. After releasing the cutting plane in the dock, the tour playback resumed.

4 EXPERIMENT

Our experiment described in this section was designed to investigate the impact of interactivity during guided tour playback for asynchronous collaboration. More specifically, we aimed to better understand how the extent of interaction afforded to an observer influences them following a tour guide presenting information in a VE. For this reason, we designed an experiment to address the following research questions:

- RQ1** How does the extent of interaction during an immersive guided tour impact an observing user’s recall?
- RQ2** How does the extent of interaction during an immersive guided tour impact an observing user’s experience?

4.1 Task

Participants were instructed to act as collaborative observers on a tour where the guide explained the inspection process of a 3D-printed object. During the tour, the guide explained different defects and their root causes, showing examples of the defects in the 3D-printed object. The participants were tasked with following the guide’s verbal and visuospatial inspection and were informed they would be tested on both the verbal and spatial content of the tour.

4.2 Measures

Various quantitative and qualitative measures were collected during the study. The Simulator Sickness Questionnaire (SSQ) was administered to measure any changes to the participant’s symptoms in oculomotor, disorientation, and nausea [27]. To assess the workload of the guided tour across the conditions, the NASA Task Load Index (NASA-TLX) was utilized [21]. The User Engagement Scale Short Form (UES-SF) was used to measure the participant’s engagement, specifically, their focused attention, perceived usability, aesthetic appeal, and reward factors from following along with the guided tour [39]. Two custom tests were administered to gauge participants’ knowledge of the verbal and spatial information covered during the guided tour. We detail the questions for each test in Table 1. Finally, an audio-recorded interview was performed after the tour to acquire open-ended feedback on the participant’s experience of their tour condition. During the audio-recorded interview, we prompted participants for a 1-10 rating as to whether they would like to experience their tour condition again with a topic they were passionate about instead. The other questions used for the interview are listed in Table 2.

4.3 Conditions

Four conditions were created encapsulating incremental combinations of our playback interactions (see Sec. 3). We detail these conditions in order of least to most interactivity during playback below.

The PASSIVE condition did not include any playback interactions for the user to perform. The tour showed the first-person point of view of the guide in a virtual display positioned 1m away from the user in an empty VE. This condition was intended to have users strictly follow the inspection from the guide’s perspective and to distinguish the implications of no viewpoint control in the VE compared to the other conditions.

The RAILS condition also did not include any playback interactions for the user to perform. However, the user had full control over their viewpoint within the VE, meaning they needed to maneuver themselves to follow the guide during the tour. For example, whenever the guide navigated upward or downward, the system automatically scaled the user, requiring them only to physically adjust their orientation to effectively follow the tour.

The NAVIGATION condition consisted of the *Multiscale Navigation* playback interactions. Playback of the tour would pause until *Navigate Down* and *Navigate Up* interactions were performed by the user to follow those of the guide.

Finally, the INTERACTIVE condition used the *Multiscale Navigation* and *Perspective Sharing* playback interactions. Playback of the tour would pause until *Navigate Down* and *Navigate Up* interactions were performed by the user to mimic the same actions of the guide. Additionally, tour playback would pause whenever the guide specified a point of interest requiring the user’s attention using the *Controller Point*, *WYSIWIS*, and *Plane Dock* interactions.

4.4 Hypotheses

We created hypotheses for each of the research questions before conducting our experiment. Regarding the influence of interaction on participant recall (**RQ1**), we hypothesized:

- H1.1** The INTERACTIVE condition will have a greater recollection of verbal information than all other conditions.
- H1.2** The INTERACTIVE condition will have a greater recollection of spatial information than all other conditions.

From a social psychological perspective of the tourist-guide relationship, we expect internal mental retention of information accompanying interaction within a given spatial environment [1, 41, 52].

In regards to the influence of interaction on the participant’s experience (**RQ2**), our hypotheses are as follows:

- H2.1** The INTERACTIVE and NAVIGATION conditions will have a higher reported engagement than the non-interactive conditions.
- H2.2** The PASSIVE condition will be reported as having a higher task workload than all other conditions.
- H2.3** The RAILS and PASSIVE conditions will be reported as having higher simulator sickness scores.

Interview Questions
Suppose that someone wanted to teach you about a topic you are passionate about using a guided tour system similar to the one you just experienced. On a scale from 1-10, with 1 being ‘not at all’ and 10 being ‘very much’, how likely would you be to take another guided tour in this format? Why?
Considering all the ways the tour tried to draw your attention to areas of interest, which interactions by the guide were helpful to you following along? Why?
Were there any interactions you would like to be able to do inside the tour that you did not experience today? If so, describe them.
Did you encounter any issues when following the tour? If so, describe them.

Table 2: Interview questions asked during the post-study.

From prior work targeting visualization and interaction within VR guided tours, we expect favorable reviews as a result of more interaction performed by a given participant [30, 40].

4.5 Design

The experiment used a between-subject design, where each participant observed a tour for only one of the conditions described in Section 4.3. Participants were randomly assigned to one of the four conditions, so each condition had 10 participants. The administration of tests described in Section 4.2 was counterbalanced, with half of the participants completing Test 1 first followed by Test 2 and vice versa. The independent variable was the extent of interaction afforded to the user during the tour. Dependent variables were the participant’s tour test scores, and subjective responses for the SSQ, UES-SF, NASA-TLX, and audio-recorded interview.

4.6 Participants

A total of 40 participants were recruited for the study using university email lists and newsletters. All were screened before participating, indicating they had normal or corrected vision, were comfortable standing for long periods, and were fluent in English. Additionally, two Likert scales between 1-6 (i.e., 1 being “not” and 6 being “very”) were completed regarding the user’s interest in and familiarity with 3D printing. For the INTERACTIVE condition, four participants were female and six were male, with ages ranging between 19 to 30 years ($M=24.70$, $SD=3.47$). Seven participants reported having experienced VR 10 or more times, one between 3-10 times, and two between 0-2 times. Likert scale values reported a mean of 2.5 for familiarity and 3.4 for interest in 3D printing. In the NAVIGATION condition, four participants were female and six were male, with ages ranging between 20 to 36 years ($M=28.10$, $SD=5.47$). Five participants reported having experienced VR 10 or more times, four between 3-10 times, and one between 0-2 times. Likert scale values reported a mean of 2.6 for familiarity and 3.9 for interest in 3D printing. For the RAILS condition, four participants were female and six were male, with ages ranging between 23 to 35 years ($M=28.00$, $SD=3.97$). Nine participants reported having experienced VR 10 or more times and one between 3-10 times. Likert scale values reported a mean of 3.4 for familiarity and 4.3 for interest in 3D printing. Finally, in the PASSIVE condition, four participants were female and six were male, with ages ranging between 23 to 35 years ($M=27.60$, $SD=4.38$). Six participants reported having experienced VR 10 or more times, two between 3-10 times, and two between 0-2 times. Likert scale values reported a mean of 2.5 for familiarity and 4.2 for interest in 3D printing.

4.7 Apparatus & Environment

Participants wore a Meta Quest 2 HWD during the study. It has a resolution of 1832x1920 per eye with a refresh rate of 90Hz and a 104-degree horizontal and 98-degree vertical Field of View (FoV)⁴. A Meta Quest 2 Touch controller, held in the participant’s specified dominant hand, was used to perform the interactions described in Section 4.3. The participant’s movement within a VE was tracked using the Meta Quest Headset Tracking system⁵. The boundary was

⁴<https://www.meta.com/quest/products/quest-2/>

⁵<https://www.meta.com/help/quest/articles/headsets-and-accessories/using-your-headset/>

created in an unobstructed area of approximately 2.5m x 2.5m and calibrated using floor markers in the physical environment before conducting each study.

Four VEs were used during the study: Setup, Training, Inspection, and Test. The Setup VE was an empty skybox environment and was strictly used to perform height calibration of the participant and facilitate swapping between the other VEs by the investigator. The Training VE was populated with props based on the interaction condition assigned to the participant. The Inspection VE had a cuboid lattice model with 1m x 1m x 1m dimensions in its center. This same environment was used for the Test VE, however, controls were added for the investigator to administer test questions. Additionally, a transparent yellow headset was shown in the Test VE, marking where the participant started the inspection tour. These VEs are shown in Fig. 2.

4.8 Procedure

Each participant completed three phases as part of the experiment: pre-study, study, and post-study. In the pre-study, the participant read and signed an informed consent document about the study and completed an SSQ. Then the participant was introduced to the HWD and had their interpupillary distance measured so that the lens positions could be set accordingly in the HWD. Finally, the participant was asked their dominant hand and shown the corresponding controller and the buttons they would need to use during the post-study and/or study according to their interaction condition. Once they indicated they were comfortable with the controller and understood how to adjust and put on the HWD, the study phase started.

In the study phase, the participant had their height calibrated inside the Setup VE to a uniform height of 1m. The investigator then placed the participant in the Training VE and started the training tour. During the training tour, the participant followed a guide explaining the interactions to be expected during the inspection tour based on their study condition and was required to perform interactions to proceed in the tour accordingly (see Section 4.3). After completing it, the investigator answered any questions regarding the interactions, marking the end of the training tour. The investigator then placed the participant in the Inspection VE. Before starting the inspection tour, the investigator gave a short description regarding the subject of the tour (i.e., the inspection of a 3D-printed object) and informed the participant that they would be tested regarding key locations shown and descriptions provided by the guide about the object from the tour. The participant then followed the inspection tour, lasting approximately eight minutes. The end of the inspection tour concluded the study phase and the participant then moved on to the final phase: the post-study.

In the post-study phase, the participant removed the HWD and completed another SSQ. After approximately three minutes passed from removing the HWD, the investigator had the participant put the HWD back on and placed the participant in the Test VE. The participant was then prompted to answer the questions from Table 1 for one of the tests, positioning themselves at the starting marker as described in Section 4.7 before being asked each question. For verbal questions, the participant answered orally and the investigator wrote down their response for later grading. For spatial questions, the participant used a wand with a red sphere at its tip attached to their controller to mark the approximate position of the prompted

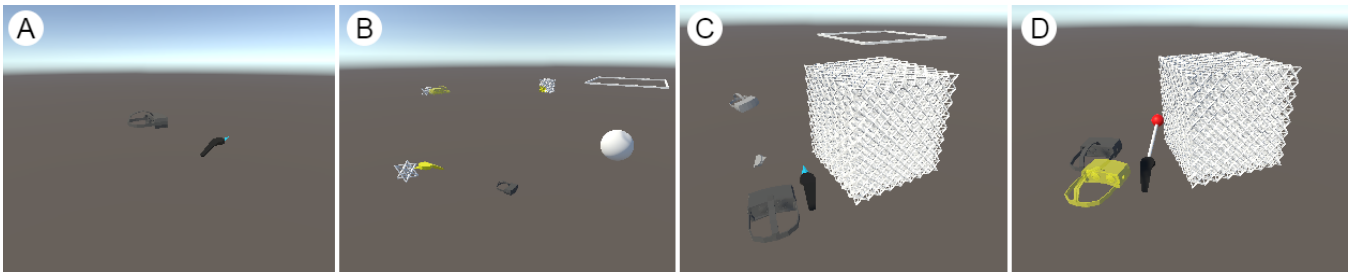


Figure 2: Exocentric views of the VEs used in the experiment. A: Setup VE showing the participant’s avatar in its empty skybox. B: Training VE showing the props used to teach all playback interactions for the INTERACTIVE condition. C: Inspection VE showing the participant’s avatar observing the guide’s avatar and the 3D-printed object. D: Test VE showing the player avatar performing a spatial marking task using the ‘marking wand’.

defect in the 3D-printed object at the highest scale, pulling the trigger once they believed the red sphere overlapped with that cell. The distance between the red sphere and the target cell was automatically logged by the system. Once they answered all the questions, the participant again removed the HWD and completed the UES-SF and NASA-TLX questionnaires. After completing the questionnaires, the investigator conducted an audio-recorded interview with the participant regarding the guided tour they experienced. Following the interview, the participant put the HWD back on and performed a buffer activity: the tutorial and demo song included in the Beat Saber - Demo⁶. This activity served to clear participants’ working memory of tour information, having them remove their focus of attention on tour content and its required sensorimotor actions by performing the same actions in an unrelated experience [26]. Once they completed the tutorial and demo, they took the remaining test in the Test VE. In this way, we hoped to measure immediate recall with the first test and medium-term retention with the second test. Finally, they completed a background questionnaire with demographic information as described in Section 4.6. The entire study took approximately 60 minutes.

5 RESULTS

We utilized various analyses to test our hypotheses using the data collected during our study. We first performed the Kruskal-Wallis tests to check for significance across dimensions for non-parametric Likert scale-based questionnaire responses. If statistical significance was found, we then performed post-hoc pairwise comparisons using Mann-Whitney U tests with Holm-Bonferroni corrections. We report this approach’s statistical significance for the SSQ, UES-SF, NASA-TLX, and condition rating results. Analysis methods for the test questions posed in Table 1 are explained in their respective subsections.

5.1 Simulator Sickness

We obtained the difference in reported simulator sickness by subtracting the pre-study SSQ from the post-study SSQ. The *Total* score was then calculated, as well as the separate sickness scores of *Disorientation*, *Nausea*, and *Oculomotor* dimensions. We did not find any significant effects between the interaction conditions on simulator sickness increase across any of these dimensions from the Kruskal-Wallis tests performed. We additionally analyzed the difference in pre- and post-study SSQ scores for each condition individually. Mann-Whitney U tests revealed significant increases in simulator sickness scores of the PASSIVE interaction condition for the *Disorientation* ($p < 0.01$), *Nausea* ($p = 0.03$), *Oculomotor* ($p < 0.01$), and *Total* ($p < 0.01$) dimensions. The INTERACTIVE condition also had significant increases in simulator sickness scores for the *Oculomotor* ($p = 0.03$) and *Total* ($p = 0.04$) dimensions.

⁶<https://www.meta.com/experiences/1758986534231171/>

These significant pairs and the mean pre- and post-study sickness scores are shown in Fig. 3a.

5.2 Engagement

Using the UES-SF data, Kruskal-Wallis results found statistically significant score differences for *Aesthetic Appeal* ($H(3) = 21.61$, $p < 0.01$), *Perceived Usability* ($H(3) = 11.90$, $p < 0.01$), and *Reward* ($H(3) = 21.61$, $p < 0.01$) dimensions. From post-hoc pairwise comparisons for *Aesthetic Appeal*, we found the PASSIVE interaction condition to be significantly less appealing than the INTERACTIVE ($p < 0.01$), NAVIGATION ($p < 0.01$), and RAILS ($p < 0.01$) conditions. For *Perceived Usability*, we found the PASSIVE interaction condition to be significantly less usable than the INTERACTIVE ($p = 0.03$), NAVIGATION ($p = 0.04$), and RAILS ($p = 0.02$) conditions. Finally, for the *Reward*, we found the PASSIVE interaction condition to be significantly less rewarding than the INTERACTIVE ($p < 0.01$), NAVIGATION ($p < 0.01$), and RAILS ($p < 0.01$) conditions. These significant pairs and mean UES-SF scale responses are shown in Fig. 3b.

5.3 Workload

Kruskal-Wallis tests across the NASA-TLX dimensions reported significant differences between interaction conditions for the *Mental Demand* ($H(3) = 8.84$, $p = 0.03$) and *Frustration* ($H(3) = 9.49$, $p = 0.02$) dimensions. After conducting post-hoc pairwise comparisons using Mann-Whitney U tests with Holm-Bonferroni corrections, no interaction condition pairs were found to be significantly different for the *Mental Demand* dimension. For the *Frustration* dimension, the PASSIVE condition was found to be significantly more frustrating than the INTERACTIVE condition ($p = 0.05$). The mean scores for the NASA-TLX are shown in Fig. 3c.

5.4 Ratings

Kruskal-Wallis tests on the reported ratings of interaction conditions indicated significant differences between them ($H(3) = 13.13$, $p < 0.01$). From post-hoc pairwise comparisons, we found that the PASSIVE interaction condition was rated significantly lower than the INTERACTIVE ($p = 0.02$), NAVIGATION ($p = 0.02$), and RAILS ($p = 0.03$) conditions. These significant pairs and mean condition ratings are shown in Fig. 3d.

5.5 Auditory Recall

We graded participant scores based on the transcribed responses to the verbal test questions posed (see Table 1). Responses were marked as either correct or incorrect based on a common grading rubric, with no partial credit being awarded. A Shapiro-Wilk test was used to check the normality of the test scores for each interaction condition, revealing that the data was not normally distributed. We then applied an Aligned Rank Transform (ART) before using a mixed Analysis of Variance (ANOVA), as the *retention time* (i.e., short and medium-term) was a within-subjects factor

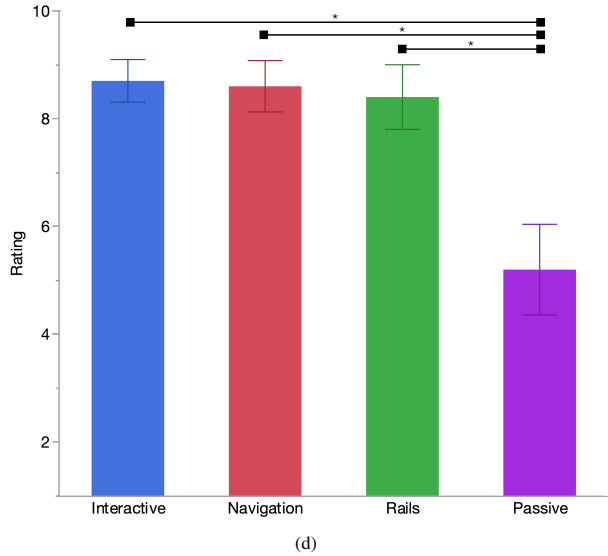
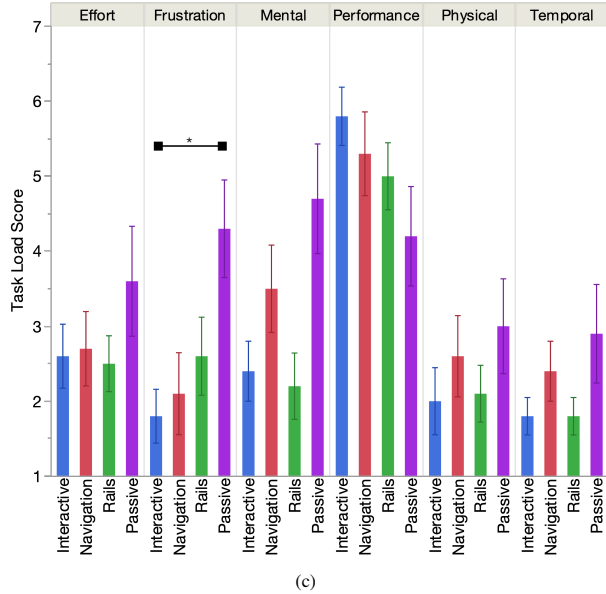
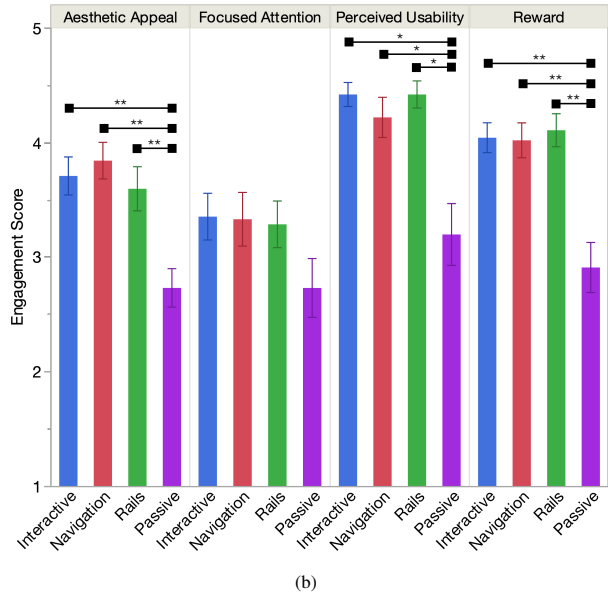
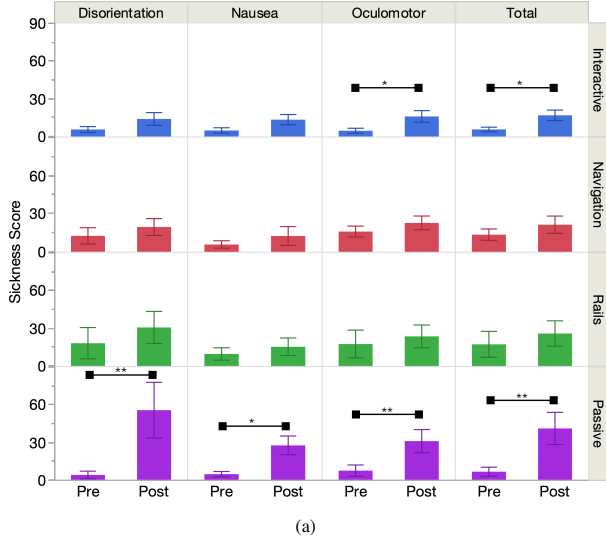


Figure 3: Bar charts of questionnaire scores. Significantly different pairs are marked with * when $p \leq 0.05$ and ** when $p \leq 0.01$. Bar chart whiskers are the $\pm S.E.$ spread of the data. (a): Mean SSQ scores from pre- and post-study. (b): Mean UES scores. (c): Mean NASA-TLX scores. (d): Mean condition ratings.

and the interactive condition a between-subjects factor [61]. The ART ANOVA indicated a significant effect of the interactive condition ($F(3, 36) = 3.16, p = 0.04$) and no effect of retention time ($F(1, 36) = 1.00, p = 0.32$) or the interaction between the interactive condition and retention time ($F(3, 36) = 0.21, p = 0.89$). Post hoc pairwise comparisons were then conducted for the interactive condition factor with the ART-C procedure, and corrected with Tukey's Honest Significance Difference (HSD) [10]. This showed that the difference in audio recall scores between the INTERACTIVE and PASSIVE conditions were statistically significant ($p = 0.04$). No other pairwise comparisons were significantly different. The mean scores across interactive conditions and the significant pairing are shown in Fig. 4a.

5.6 Spatial Recall

A Shapiro-Wilk test was first used to check the normality of reported defect distances for each interaction condition, revealing that

the data was not normally distributed. We then applied an ART before using a mixed ANOVA for the within-subjects factor of retention time and the between-subjects factor of interactive condition. The ART ANOVA indicated a significant effect of the interactive condition ($F(3, 36) = 4.18, p = 0.01$) and no effect of retention time ($F(1, 36) = 1.30, p = 0.26$) or the interaction between the interactive condition and retention time ($F(3, 36) = 1.72, p = 0.18$). Post hoc pairwise comparisons were then conducted for the interactive condition factor with the ART-C procedure, and corrected with Tukey's HSD. This showed that the difference in spatial recall distances between the INTERACTIVE and PASSIVE conditions were statistically significant ($p = 0.04$), as well as the RAILS and PASSIVE conditions ($p = 0.02$). No other pairwise comparisons were significantly different. The mean reported distances and significance pairings are shown in Fig. 4b.

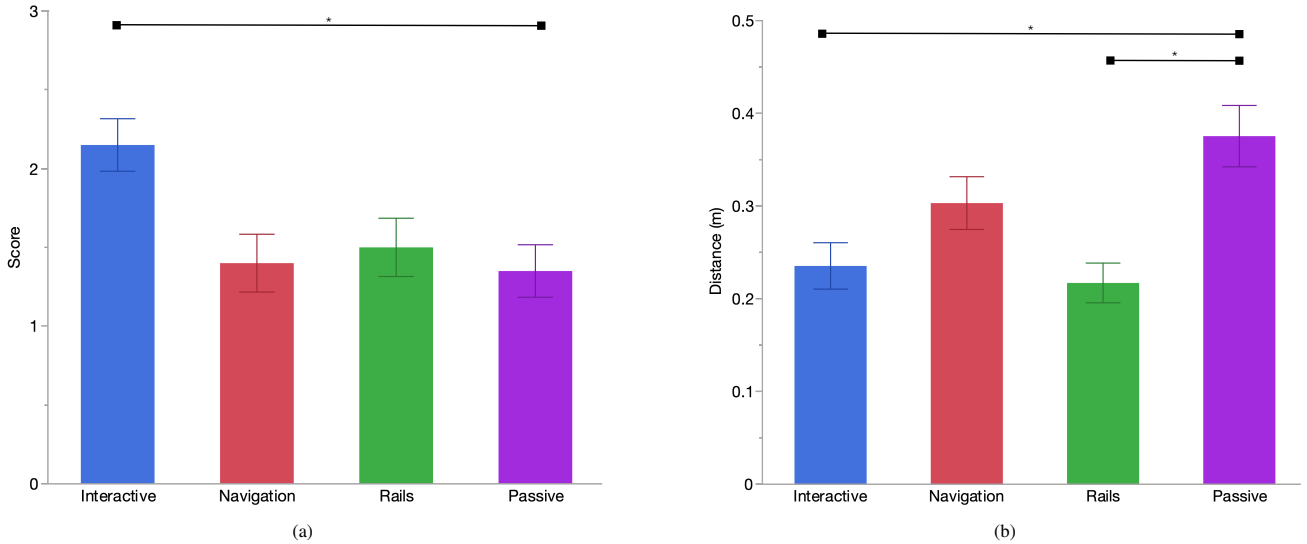


Figure 4: Bar charts of collected test responses for verbal and spatial questions from Table 1. (a): Verbal scores. (b): Distances of participant-reported defect location from the actual defect location.

6 DISCUSSION

6.1 Recall

To assess the participant’s recollection of information covered during the guided tour, we created a set of test questions (i.e., Table 1) regarding the locations of defects within the 3D object and dialogue spoken by the guide throughout the tour. These tests were administered immediately following the tour and after a buffer activity, as described in Sec. 4.8. This was intended to capture their retention of the tour in the short term and following a verbal and visuospatial buffer [26]. We used the test responses and open-interview transcripts to answer **RQ1** and test the corresponding hypotheses: *H1.1* and *H1.2*.

We hypothesized that participants who experienced the tour with the INTERACTIVE condition would have a greater recollection of verbal information than the other conditions (*H1.1*). Our results partially support *H1.1*, as those who participated in the INTERACTIVE tour achieved significantly higher scores compared to those in the PASSIVE condition. However, despite INTERACTIVE achieving the highest average score, it was not found to be significantly different than NAVIGATION and RAILS. Although all participants listened to the same tour narrative from the guide, this result suggests that interaction during the embodied inspection proved crucial in retaining the auditory information presented. We believe this follows the view in sensorimotor enactivism that one’s perception relies on the active exploration of an environment, consisting of movements and sensory states in that environment, where the participant’s auditory perception was improved by the additional playback actions [46]. The *Controller Point* action was described by several participants in the INTERACTIVE condition as helping them direct their attention to and remember areas of interest. *P5* mentioned “the guide telling me to look at the point... I feel like I still might have missed it unless I actually went into that position and put my hand exactly where it was pointing”, *P17* stated that they “liked when the guide made me point... it’s clicking to the point and being like ‘Yeah, I see it’... it’s like we’re in the same spot”, and *P37* described that when “the tour guide sees something you don’t quite see yet, they point at it... you kind of get that ‘aha’ moment of like, ohh, that’s what we’re looking at.” From the other tour conditions, participants described their desire to further interact with the environment influencing their retention of information. *P6* in the NAVIGATION condition mentioned “if I was able to interact more... I think I would have been able to retain the information

a little better”, *P31* of the RAILS condition described preferring “to have some time for myself to navigate through the object after [the guide] explained... I could test and review the object for myself or see that information”, and *P12* in the PASSIVE condition stated “I prefer being more hands-on with things and I wish I had control as far as like when I could proceed and I almost wish there was a pause function so you could, like, stop and look instead of here’s this thing, here’s that thing.”

We additionally hypothesized that participants who experienced the tour with the INTERACTIVE condition would have a greater recollection of spatial information than all other conditions (*H1.2*). Our results partially support *H1.2*, as those who participated in the INTERACTIVE tour achieved significantly better estimates of the approximate locations of defects compared to those in the PASSIVE condition. However, the same statistical significance was found for the RAILS condition when compared to the PASSIVE condition. This is similar to a comparison study examining viewpoint control, where the participants who had independent control over their viewpoint achieved higher test scores than those with a fixed viewpoint [33]. This implies a greater influence of viewpoint in visuospatial recall, rather than our anticipated grounding that the participant benefits from body movements required by our interactions to increase environment understanding [46]. A missing pairing of interest was the significance between the NAVIGATION and PASSIVE conditions, where we would similarly expect the independent control of viewpoint resulting in better recollection of the approximate positions of defects. We attribute this to our guided tour system, which allowed participants to perform the *Navigation Down* and *Navigation Up* at any point during the INTERACTIVE and NAVIGATION conditions, resulting in many participants preemptively navigating across scales, requiring them to navigate back to the guide to resume playback. A few participants described this in the post-study interview, such as *P21* from the INTERACTIVE condition stating that their “only issue would be going too fast. Hearing like ‘let’s zoom in’ and then doing it before I heard the pause sound effect” and *P42* in the NAVIGATION condition describing “I got too curious at some points and when the guy was guiding me to a different point of interest for the struts I was like ‘I really want to go see it’ and I would. I did not wait and went rogue at some points.”

Overall, the tours using the INTERACTIVE condition had a greater auditory recall and better spatial recall than the PASSIVE condition. Although the INTERACTIVE condition did not show

significantly better recall compared to the other independent viewpoint conditions, it produced the highest performance, which could potentially be enhanced with adjustments to the playback system. The evidence, therefore, suggests that additional interactivity beyond simple viewpoint control can help increase observers' recall of guided tour content.

6.2 User Experience

To assess the participants' experience from following the guided tour, we leveraged existing research questionnaires including the SSQ, UES-SF, and NASA-TLX, as well as post-study ratings from 1-10 if participants would participate in another tour in the same format they experienced. These questionnaires were completed by the participant throughout the study as documented in the procedure. We used the responses and open-interview transcripts to answer **RQ2** and test the affiliated hypotheses: *H2.1*, *H2.2*, and *H2.3*.

We hypothesized that the INTERACTIVE and NAVIGATION conditions would have higher reported engagement than the RAILS and PASSIVE conditions (*H2.1*). Based on the UES-SF response data, our results partially support *H2.1* with an interesting exception. The RAILS condition, although not having controller-based interaction, was reported as being approximately as engaging as the INTERACTIVE and NAVIGATION conditions, with significance existing from higher reported scores across *Aesthetic Appeal*, *Perceived Usability*, and *Reward* dimensions of the UES-SF than the PASSIVE condition. Reviewing open-interview transcripts from participants who completed the PASSIVE condition, the reasons for these lower scores could be found. For instance, lack of engagement using VR was an issue mentioned by several participants, such as *P12* who stated "it was almost kind of like you were doing it in VR and then you were kind of being drug around", *P36* describing "for me it's about control. I didn't have that during the video... didn't have my own movement and head movements", and *P44* saying "watching a video in VR kind of defeats the purpose... since it's just a video, I'm not interacting with anything."

We also hypothesized that the PASSIVE condition would be reported as having a higher task workload than all other conditions (*H2.2*). From the NASA-TLX response data, our results do not support *H2.2*. While the PASSIVE condition did have worse task load scores across all dimensions of the NASA-TLX, only the *Frustration* dimension had a significant pairing with the INTERACTION condition. Still, participants explicitly mentioned issues with task load when discussing the PASSIVE condition in the post-study interview. Regarding the *Temporal* load, *P40* described feeling the tour "was fast... a lot of information in a very short time" and *P28* stated that the tour was "a bit too demanding and I think it was a bit fast-paced, so kind of keeping track was a bit difficult." Additional dimensions mentioned were the *Mental* load, where *P36* stated that "to memorize someone else's movement and what they were saying and what I was seeing... that created a lot of mental load", and the *Frustration*, with *P12* remarking that the tour was simply "really hard to follow."

Finally, we hypothesized that the RAILS and PASSIVE conditions would have higher reported simulator sickness scores than the INTERACTIVE and NAVIGATION conditions (*H2.3*). From the SSQ response data, our results do not support *H2.3*. Although the PASSIVE condition did have significant increases in simulator sickness scores between pre- and post-study measurements across all dimensions, the RAILS condition did not have any significant increases. Furthermore, the INTERACTIVE condition reported having significant increases in simulator sickness for the *Oculomotor* and *Total* dimensions. We attribute this to *WYSIWIS*, where participants reported discomfort when performing the playback interaction. *P13* stated that they "didn't love [WYSIWIS]... having the transparent yellow headset and having to look... my view was being stuck on the headset", *P17* described that the "[WYSIWIS] blocked my

view and it was weird", and *P25* said "the headset one... is a little disorienting."

Overall, the tour using the PASSIVE condition was found to have a worse user experience across multiple engagement, task workload, and simulator sickness measures. Along with their final ratings, participants found the PASSIVE condition to be limited in its usability and subsequent engagement. Conversely, the other conditions were comparably rated across these measures, making them all viable means to conduct tours from the user's perspective. We attribute this difference to the distinct design of the PASSIVE condition, in which participants did not control the tour's viewpoint, experiencing it instead from a third-person perspective. This implies a greater influence of viewpoint on the user's experience as opposed to strictly the extent of interaction during the guided tour we expected.

6.3 Limitations & Future Work

Our work had several limitations. First, the subject matter of 3D printing was not uniformly understood or of interest to the participants across conditions. This could have affected the recall and experience of the participant. Future work could consist of a more generalized topic outside of advanced manufacturing to be exciting and more universally understood, such as sports coach-player mentoring [31], medicine including anatomy and immunology [64, 62], or general education topics, such as history and cultural heritage [45]. Secondly, the sample size of our study was relatively small due to the between-subjects design. A larger sample size may have resulted in us finding significance in cases where values were trending towards a significant difference, such as the NASA-TLX dimensions. Future studies should recruit more participants to thoroughly analyze the recall outcomes and user experience. Alternatively, the same participant quantity could be used by omitting a condition for comparison, providing more data to analyze between the remaining conditions. Additional tours within the VE for unique inspection objects could also provide a means to design a within-subjects experiment. Thirdly, the asynchronous collaboration was focused on a confederate recording performed by the investigator and viewed by the observing participant. A future study could focus on having a preceding participant perform a recording for the following participant to observe, such that insights regarding the recording production process can be considered. Lastly, the 3D printed object used for inspection was homogenous due to the unit cell models it was comprised of. Future studies could consider using an object with distinct differences in its geometry or markers to provide reference points for determining relative locations of points of interest covered during a tour.

7 CONCLUSION

In this work, we document our design of a guided tour system to facilitate asynchronous collaboration for AM inspections. We describe the use of guide mimicry by the observing user as a means to progress the playback of a guided tour. Our results from a user study investigating the extent of using these mimicking actions suggest that interactivity positively influences auditory and spatial recall and that the extent of viewpoint control helps improve the user's experience when following a guided tour. For future work, we plan to investigate the use of guided tours in more generalized collaborative topics.

ACKNOWLEDGMENTS

This work was supported by the US DOE LLNL-LDRD 23-SI-003 and was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-CONF-869200).

REFERENCES

- [1] M. Argyle et al. *Social situations*. Cambridge University Press, 1981. 4
- [2] B. Bell et al. View management for virtual and augmented reality. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology*, UIST '01, p. 101–110. Association for Computing Machinery, New York, NY, USA, 2001. doi: 10.1145/502348.502363 2
- [3] K. Best. Making museum tours better: understanding what a guided tour really is and what a tour guide really does. *Museum Management and Curatorship*, 27(1):35–52, 2012. doi: 10.1080/09647775.2012.644695 2
- [4] D. A. Bowman et al. Information-rich virtual environments: theory, tools, and research agenda. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '03, p. 81–90. Association for Computing Machinery, New York, NY, USA, 2003. doi: 10.1145/1008653.1008669 2
- [5] V. Chheang et al. A virtual environment for collaborative inspection in additive manufacturing. In *Extended Abstracts of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI EA '24. Association for Computing Machinery, New York, NY, USA, 2024. doi: 10.1145/3613905.3650730 2
- [6] H. Cho et al. Realityreplay: Detecting and replaying temporal changes in situ using mixed reality. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 7(3), Sept. 2023. doi: 10.1145/3610888 3
- [7] K. Chow et al. Challenges and design considerations for multimodal asynchronous collaboration in vr. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), nov 2019. doi: 10.1145/3359142 3
- [8] E. F. Churchill et al. Collaborative virtual environments: an introductory review of issues and systems. *virtual reality*, 3:3–15, 1998. 1
- [9] D. Clergeaud et al. Design of an annotation system for taking notes in virtual reality. In *2017 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, pp. 1–4, 2017. doi: 10.1109/3DTV.2017.8280398 2
- [10] L. A. Elkin et al. An aligned rank transform procedure for multi-factor contrast tests. In *The 34th Annual ACM Symposium on User Interface Software and Technology*, UIST '21, p. 754–768. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474784 7
- [11] B. Ens et al. Revisiting collaboration through mixed reality: The evolution of groupware. *International Journal of Human-Computer Studies*, 131:81–98, 2019. 50 years of the International Journal of Human-Computer Studies. Reflections on the past, present and future of human-centred technologies. doi: 10.1016/j.ijhcs.2019.05.011 1
- [12] A. R. Fender et al. Causality-preserving asynchronous reality. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, CHI '22. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3491102.3501836 3
- [13] E. Frécon et al. Building distributed virtual environments to support collaborative work. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology*, VRST '98, p. 105–113. Association for Computing Machinery, New York, NY, USA, 1998. doi: 10.1145/293701.293715 2
- [14] T. A. Galyean. Guided navigation of virtual environments. In *Proceedings of the 1995 Symposium on Interactive 3D Graphics*, I3D '95, p. 103–ff. Association for Computing Machinery, New York, NY, USA, 1995. doi: 10.1145/199404.199421 2
- [15] I. García-Pereira et al. Cross-device augmented reality annotations method for asynchronous collaboration in unprepared environments. *Information*, 12(12), 2021. doi: 10.3390/info12120519 3
- [16] R. K. Ghamandi et al. What and how together: A taxonomy on 30 years of collaborative human-centered xr tasks. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 322–335, 2023. doi: 10.1109/ISMAR59233.2023.00047 1
- [17] A. Giovannelli et al. Gestures vs. emojis: Comparing non-verbal reaction visualizations for immersive collaboration. *IEEE Transactions on Visualization and Computer Graphics*, 29(11):4772–4781, 2023. doi: 10.1109/TVCG.2023.3320254 1
- [18] T. J. Guerreiro et al. Beyond post-it: Structured multimedia annotations for collaborative ves. In *ICAT-EGVE*, pp. 55–62, 2014. 2
- [19] F. A. Hansen. Ubiquitous annotation systems: technologies and challenges. In *Proceedings of the Seventeenth Conference on Hypertext and Hypermedia*, HYPERTEXT '06, p. 121–132. Association for Computing Machinery, New York, NY, USA, 2006. doi: 10.1145/1149941.1149967 2
- [20] R. Harmon et al. The virtual annotation system. In *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, pp. 239–245, 1996. doi: 10.1109/VRRAIS.1996.490533 2
- [21] S. G. Hart. Nasa-task load index (nasa-tlx); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006. doi: 10.1177/154193120605000909 4
- [22] L. Hong et al. Virtual voyage: interactive navigation in the human colon. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '97, p. 27–34. ACM Press/Addison-Wesley Publishing Co., USA, 1997. doi: 10.1145/258734.258750 2
- [23] A. Irlitti et al. Tangible interaction techniques to support asynchronous collaboration. In *2013 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 1–6, 2013. doi: 10.1109/ISMAR.2013.6671840 2
- [24] A. Irlitti et al. Challenges for asynchronous collaboration in augmented reality. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp. 31–35, 2016. doi: 10.1109/ISMAR-Adjunct.2016.0032 1, 2
- [25] R. Johansen. Groupwise and collaborative systems—a big picture view. In *1989 IEEE Global Telecommunications Conference and Exhibition 'Communications Technology for the 1990s and Beyond'*, pp. 1217–1220, 1989. doi: 10.1109/GLOCOM.1989.64148 2
- [26] J. Jonides et al. The mind and brain of short-term memory. *Annu. Rev. Psychol.*, 59:193–224, 2008. 6, 8
- [27] R. S. Kennedy et al. Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220, 1993. doi: 10.1207/s15327108ijap0303_3 4
- [28] P. Klacansky et al. Virtual inspection of additively manufactured parts. In *2022 IEEE 15th Pacific Visualization Symposium (PacificVis)*, pp. 81–90, 2022. doi: 10.1109/PacificVis53943.2022.00017 3
- [29] G. A. Lee et al. Enhancing first-person view task instruction videos with augmented reality cues. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 498–508, 2020. doi: 10.1109/ISMAR50242.2020.00078 3
- [30] J. Letellier et al. Visualization and interaction techniques in virtual reality for guided tours. In *2019 10th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS)*, vol. 2, pp. 1041–1045, 2019. doi: 10.1109/IDAACS.2019.8924392 5
- [31] T. Lin et al. Towards an understanding of situated ar visualization for basketball free-throw training. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3411764.3445649 9
- [32] L. Lisle et al. Sensemaking strategies with immersive space to think. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pp. 529–537, 2021. doi: 10.1109/VR50410.2021.00077 2
- [33] R. Lovreglio et al. Comparing the effectiveness of fire extinguisher virtual reality and video training. *Virtual Reality*, 25(1):133–145, 2021. 8
- [34] B. Marques et al. Remote asynchronous collaboration in maintenance scenarios using augmented reality and annotations. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 567–568, 2021. doi: 10.1109/VRW52623.2021.00166 1, 3
- [35] B. Marques et al. Which notification is better? comparing visual, audio and tactile cues for asynchronous mixed reality (mr) remote collaboration: A user study. In *Proceedings of the 21st International Conference on Mobile and Ubiquitous Multimedia*, MUM '22, p. 276–278. Association for Computing Machinery, New York, NY, USA, 2022. doi: 10.1145/3568444.3570587 1, 3

- [36] A. Mayer et al. Asynchronous manual work in mixed reality remote collaboration. In *Extended Reality*, pp. 17–33. Springer Nature Switzerland, Cham, 2022. 3
- [37] A. Mayer et al. *Collaborative Work Enabled by Immersive Environments*, pp. 87–117. Springer International Publishing, Cham, 2023. doi: 10.1007/978-3-031-26490-0_6 1, 2, 3
- [38] P. Milgram et al. Augmented reality: a class of displays on the reality-virtuality continuum. In H. Das, ed., *Telem manipulator and Telepresence Technologies*, vol. 2351, pp. 282–292. International Society for Optics and Photonics, SPIE, 1995. doi: 10.1117/12.197321 3
- [39] H. L. O'Brien et al. A practical approach to measuring user engagement with the refined user engagement scale (ues) and new ues short form. *International Journal of Human-Computer Studies*, 112:28–39, 2018. doi: 10.1016/j.ijhcs.2018.01.004 4
- [40] R. Pausch et al. Disney's aladdin: first steps toward storytelling in virtual reality. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '96*, p. 193–203. Association for Computing Machinery, New York, NY, USA, 1996. doi: 10.1145/237170.237257 2, 5
- [41] P. L. Pearce. Tourist-guide interaction. *Annals of Tourism Research*, 11(1):129–146, 1984. 4
- [42] S. Pick et al. Automated positioning of annotations in immersive virtual environments. In *Proceedings of the 16th Eurographics Conference on Virtual Environments & Second Joint Virtual Reality, EGVE - JVRC'10*, p. 1–8. Eurographics Association, Goslar, DEU, 2010. 2
- [43] C. Pidel et al. Collaboration in virtual and augmented reality: A systematic overview. In L. T. De Paolis and P. Bourdot, eds., *Augmented Reality, Virtual Reality, and Computer Graphics*, pp. 141–156. Springer International Publishing, Cham, 2020. 1, 2
- [44] W. Ribarsky et al. Visualization and analysis using virtual reality. *IEEE Computer Graphics and Applications*, 14(1):10–12, 1994. doi: 10.1109/38.250911 2
- [45] M. Roussou. Learning by doing and learning through play: an exploration of interactivity in virtual environments for children. *Comput. Entertain.*, 2(1):10, Jan. 2004. doi: 10.1145/973801.973818 9
- [46] L. Shapiro et al. Embodied Cognition. In E. N. Zalta and U. Nodelman, eds., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2024 ed., 2024. 8
- [47] A. Steed et al. Collaboration in immersive and non-immersive virtual environments. In *Immersed in Media*, pp. 263–282. Springer, 2015. 1
- [48] M. Stefik et al. Wysiwis revised: early experiences with multiuser interfaces. *ACM Trans. Inf. Syst.*, 5(2):147–167, apr 1987. doi: 10.1145/27636.28056 3
- [49] I. A. Tahmid et al. Evaluating the benefits of explicit and semi-automated clusters for immersive sensemaking. In *2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 479–488, 2022. doi: 10.1109/ISMAR55827.2022.00064 2
- [50] I. A. Tahmid et al. Colt: Enhancing collaborative literature review tasks with synchronous and asynchronous awareness across the reality-virtuality continuum. In *2023 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, pp. 831–836, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00183 2
- [51] J. Tam et al. A framework for asynchronous change awareness in collaborative documents and workspaces. *International Journal of Human-Computer Studies*, 64(7):583–598, 2006. Theoretical and empirical advances in groupware research. doi: 10.1016/j.ijhcs.2006.02.004 1
- [52] P. Thagard. Cognitive Science. In E. N. Zalta and U. Nodelman, eds., *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2023 ed., 2023. 4
- [53] J. Thomas et al. A communication-focused framework for understanding immersive collaboration experiences. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 301–304, 2023. doi: 10.1109/VRW58643.2023.00070 1
- [54] R. J. W. Tijssen et al. Research collaboration and the expanding science grid: Measuring globalization processes worldwide, 2012. doi: 10.48550/ARXIV.1203.4194 1
- [55] E. E. Tsiropoulou et al. Quality of experience-based museum touring: a human in the loop approach. *Social Network Analysis and Mining*, 7(1):33, Jul 2017. doi: 10.1007/s13278-017-0453-2 2
- [56] J. C. Verlinden et al. *Virtual annotation: Verbal communication in virtual reality*. Delft University of Technology, Faculty of Technical Mathematics and Informatics, 1993. 2
- [57] C. Y. Wang et al. Vr-replay: Capturing and replaying avatars in vr for asynchronous 3d collaborative design. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pp. 1215–1216, 2019. doi: 10.1109/VR.2019.8797789 3
- [58] T. Wang et al. Capturar: An augmented reality tool for authoring human-involved context-aware applications. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology, UIST '20*, p. 328–341. Association for Computing Machinery, New York, NY, USA, 2020. doi: 10.1145/3379337.3415815 3
- [59] T. Wang et al. Gesturar: An authoring system for creating freehand interactive augmented reality applications. In *The 34th Annual ACM Symposium on User Interface Software and Technology, UIST '21*, p. 552–567. Association for Computing Machinery, New York, NY, USA, 2021. doi: 10.1145/3472749.3474769 3
- [60] J. Wither et al. Annotation in outdoor augmented reality. *Computers & Graphics*, 33(6):679–689, 2009. doi: 10.1016/j.cag.2009.06.001 2
- [61] J. O. Wobbrock et al. The aligned rank transform for nonparametric factorial analyses using only anova procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11*, p. 143–146. Association for Computing Machinery, New York, NY, USA, 2011. doi: 10.1145/1978942.1978963 7
- [62] L. Zhang et al. Exploring effects of interactivity on learning with interactive storytelling in immersive virtual reality. In *2019 11th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, pp. 1–8, 2019. doi: 10.1109/VS-Games.2019.8864531 9
- [63] L. Zhang et al. Vrgit: A version control system for collaborative content creation in virtual reality. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI '23*. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3544548.3581136 1
- [64] J. Zhao et al. The effectiveness of virtual reality-based technology on anatomy teaching: a meta-analysis of randomized controlled studies. *BMC medical education*, 20:1–10, 2020. 9