MR-MDEs: Exploring the Integration of Mixed Reality into Multi-display Environments

GRIFFIN J. HURT, University of Pittsburgh, USA
TALHA KHAN, University of Pittsburgh, USA and Grand Valley State University, USA
NICOLÁS MATHEO KASS, University of Pittsburgh, USA
ANTHONY TANG, University of Pittsburgh, USA
EDWARD ANDREWS, University of Pittsburgh, USA
JACOB BIEHL, University of Pittsburgh, USA

Fig. 1. The mixed reality and large display (*MR+LD*) experimental configuration for our study. Participants assembled two Lego sets using instructions provided by an multi-display environment (MDE) and were forced to switch between them frequently and randomly. To simulate a public, shared information resource, participants are provided with instructions on the large display in front of them, as well as distractors that appear visually similar to their current task. Simultaneously, the user is provided with an MR headset that provides only the instructions for the current two sets, integrating both virtual and physical displays into a unified MDE. Lego assembly was used as the experimental task to provide a point of comparison to prior work that has studied the use of MR in task guidance [3, 62].

Authors' Contact Information: Griffin J. Hurt, griffhurt@pitt.edu, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; Talha Khan, khanm4@gvsu.edu, University of Pittsburgh, Pittsburgh, Pennsylvania, USA and Grand Valley State University, Allendale, Michigan, USA; Nicolás Matheo Kass, kassn@upmc.edu, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; Anthony Tang, Tang.Anthony@medstudent.pitt.edu, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; Edward Andrews, andrewse2@upmc.edu, University of Pittsburgh, Pittsburgh, Pennsylvania, USA; Jacob Biehl, biehl@pitt.edu, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.

ACM 2573-0142/2025/12-ARTISS017
https://doi.org/10.1145/3773074

Multi-display environments (MDEs) are applicable to both everyday and specialized tasks like cooking, appliance repair, surgery, and more. In these settings, displays are often affixed in a manner that prevent reorientation, forcing users to split their attention between multiple visual information sources. Mixed reality (MR) has the potential to transform these spaces by presenting information through virtual interfaces that are not limited by physical constraints. While MR has been explored for single-task work, its role in multi-task, information-dense environments remains relatively unexplored. Our work bridges this gap by investigating the impact of different display modalities (large screens, tablets, and MR) on performance and perception in these environments. Our study's findings demonstrate the capability for MR to integrate into these spaces, extending traditional display technology with no impact to performance, cognitive load, or situational awareness. The study also further illustrates the nuanced relationship between performance and preference in tools used to guide task work. We provide insights toward the eventual authentic integration of MR in MDEs.

CCS Concepts: • **Human-centered computing → Mixed / augmented reality**; **Displays and imagers**.

Additional Key Words and Phrases: Segmented displays, virtual displays, display replacement, multi-display environments.

ACM Reference Format:
Griffin J. Hurt, Talha Khan, Nicolás Matheo Kass, Anthony Tang, Edward Andrews, and Jacob Biehl. 2025. MR-MDEs: Exploring the Integration of Mixed Reality into Multi-display Environments. *Proc. ACM Hum.-Comput. Interact.* 9, 8, Article ISS017 (December 2025), 22 pages. https://doi.org/10.1145/3773074

1 Introduction

Mixed reality (MR) allows users to place and interact with virtual content freely within the physical world. This capability of MR is fundamentally different from traditional display modalities (e.g., tablets and cellphones) where interaction is usually confined to the physical form factor of the devices [37]. Using MR, engineers can prototype complex 3D architecture models in an egocentric manner [15, 70, 72], and surgeons can overlay medical imaging directly on patients' anatomy [5, 29, 30, 40]. These benefits of MR extend to collaborative settings as well. MR enables geographically distributed workers to interact from thousands of miles away as if they were physically present together [21, 39, 64, 71, 75]. These examples highlight how MR extends the interaction capabilities of conventional display devices, offering a more dynamic and immersive interaction experience, while enabling new approaches to information-mediated physical tasks.

Research has shown that MR provides benefits in supporting physical tasks [10, 38, 42, 54, 55, 62, 65]. These prior studies have shown MR can reduce cognitive load, task errors, and task completion times [62, 65]. For example, Tang *et al.* [62] showed that AR instructions resulted in reduced error rates and lower cognitive load levels for a 3D object assembly task. Blattgerste *et al.* [8] reported comparable findings for cognitive load reduction for a similar task. Despite the proven value and potential of MR, there have been few studies that have investigated the integration of MR into existing multi-display, multi-task, interaction-limited workspaces.

A poignant, albeit specialized, example of this type of workspace is an operating room. Within this environment, medical staff are unable to interact with any non-sterile physical devices, including displays. These interaction limitations also exist in more common contexts, like preparing a large family dinner or diagnosing and fixing appliances. During these activities, the user's hands are often occupied or dirty which may prevent them from interacting with physical displays or traditional input devices such as a mouse and keyboard. Further, the spatial constraints of the environment will likely restrict the effective placement and interactive capabilities of display technology [36, 57, 66]. With the use of MR in these spaces, traditional information sources (i.e., displays) can be converted into virtual information sources. These virtual displays give users the ability to place content in locations and positions not afforded by physical displays. This capability allows the user to reduce splitting their attention, which has known negative impacts on task performance (e.g., longer task times and higher error rates) [14, 60] and learning [52] (e.g., higher cognitive load) and improve overall information ergonomics [46, 74].

In this work, we seek to understand how well MR, with its current capabilities, can be integrated into an MDE. We specifically investigated two research questions:

- **RQ1**: How does incorporating MR as a display modality in an MDE affect the user's performance, cognitive load, and situational awareness compared to traditional MDE display configurations?
- **RQ2**: How do users perceive contemporary MR's role and integrative capacity within MDEs?

We conducted a comparative study to assess the effects of three different multi-display modalities (baseline large display only, conventional tablet + large display, and mixed reality + large display) on task performance, cognitive load, and situational awareness while switching between two information-mediated, fine-motor tasks. Our study revealed several key findings. We found that MR maintains comparable task performance to other display modalities for simultaneous fine-motor tasks akin to those performed in traditional MDEs. We note a trend between modality preference and performance, which reinforces the findings of prior work that users' acceptance of a technology and their performance are related [11, 67]. We situate our study findings within the current form factor and UX design of MR hardware and position its current readiness to be integrated into exisiting MDEs. We discuss the impact of the current form factor and UX design of MR on its capability to be integrated into existing MDEs, as well as how this affects user preferences. Using these findings, we synthesize new design insights to guide the continued development of MR-based tools and applications for use in these environments.

2 Related Work

2.1 Multi-Display Environments and Ubiquitous Analytics

In this work, we specifically investigate the role of MR within multi-display environments (MDEs). MDEs, sometimes called multi-surface environments, are described in the literature as "computing environments where interaction spans multiple input and output devices and can be performed by several users simultaneously" [19, 20]. MDEs span a variety of workspaces, including airplane cockpits [43, 69], living rooms [58, 69], kitchens [45], and operating rooms [7]. Broadly, MDEs have been studied in the context of collaborative task support [6, 26, 28, 47, 48], although some work also investigates its use in single-user tasks [24]. Prior work has investigated MR's role in these environments, primarily in the form of Computer Assisted Virtual Environments, or CAVEs, finding that increasing immersion within MDEs can improve task performance [18, 41, 53].

The integration of different display modalities specifically to understand and view data has been previously studied in the literature under the concept of "ubiquitous analytics" (*ubilytics*), a specialized example of ubiquitous computing [73]. The term was first coined by Elmqvist and Irani to describe multiple networked devices in a shared space that each provide specific visualization and interaction affordances, allowing knowledge workers to better understand massive, hetereogenous data [16]. In their seminal work, they mention the possibility of integrating MR into these environments; however, capable hardware was not available at the time of their work to meaningfully test integration.

Since then, systems such as *Vistribute* [23],*Webstrates* [35], *Vistrates* [1], and *Wizualization* [4] have been developed to realize ubilytics workflows (with some even integrating MR), but these works have not yet studied the performance and perception differences across competing modalities.

2.2 Mixed Reality in Task-Guided Work

Prior work has extensively explored the use of MR in task-guided work. Daling and Schlittmeier analyzed 24 articles that detailed the use of MR in manual assembly tasks [10]. They found consistent evidence that MR's use in task guidance enabled greater flexibility in task pacing and greater understanding of task related instructions. Kumaravel et al. designed a system that allows a user to make 2D annotations with a tablet for another user to view in a virtual reality environment [65]. To test their system, they asked participants to assemble a robot in virtual reality using predefined blocks, with instructions provided by another participant using the tablet. They found that using the system resulted in significantly higher task success and significantly lower task completion time, task load, and error rate. These findings align with other studies in the broader field that show MR can improve how people perform task-guided work [12, 38, 42, 54, 55].

2.2.1 Lego Assembly. Many works in the literature have used Lego assembly as an experimental task for evaluating the effectiveness of MR as a guidance tool. Khuong *et al.* investigated the effects of different MR visualizations and error detection for Lego assembly [34]. They found that a side-by-side representation, with the instructions shown on a virtual model rendered next to the physical model, was most preferred and significantly decreased task completion time compared to a partial wireframe visualization. Tang *et al.* found that cognitive load was significantly lower when participants were presented with instructions in MR compared to a large monitor [62]; however, they found no significant difference in task performance between modalities. Blattgerste *et al.* [8] conducted a study to compare the effects of MR instructions to paper instructions for a manual Lego assembly task. Their results showed that users made fewer errors in Lego set construction when using MR, although task completion time was greater compared to paper instructions. Interestingly, Blattgerste *et al.* did not encounter the same decrease in cognitive load seen in Tang *et al.*. This demonstrates that the environment in which MR is used to provide task guidance can have significant effects on performance and further reinforces the need to study task-guidance with MR in MDEs.

2.2.2 Commercial Solutions. Task-guidance with MR also has a commercial footprint. Emergent and particularly powerful use cases of these commercial solutions are in the work domains of surgery, engineering, architecture, and 3D environment design. A prominent example in surgery is Medivis *SurgicalAR*[1], a platform which enables surgeons to overlay medical scans (e.g., X-ray/CT) on the patient's anatomy to facilitate surgical tool navigation inside human bodies. In the engineering realm, *TheoremXR*[2] extends CAD software, and allows experienced technicians to create mechanical task instructions that trainees can visualize in-situ with MR.

Both research and commercial MR solutions for task-guidance have been implemented and studied for single-task scenarios utilizing a single information source that is either retrieved from a display or presented in spatial context using MR. Our work seeks to understand how MR can be integrated into existing display-rich environments to perform individual tasks.

2.3 Mixed Reality Displays

The use of MR as a replacement to traditional planar displays has been explored by the community. Broadly, this work has shown evidence of equivalence in task performance while user preference is quite varied. Khan *et al.* explored the capacity for MR to replace planar displays in the operating room at varying levels of latency by asking participants to lay lines of suture [33]. They found no significant difference in performance and cognitive load between the monitor and MR modalities at

[1]https://www.medivis.com/surgical-ar
[2]https://www.theorem.com/extended-reality

lower latency levels. Pavanatto *et al.* evaluated the effectiveness of holographic displays in MR as a substitute. They asked participants to perform a grading task using a purely virtual, hybrid, and a purely physical setup [50]. They found that participants were significantly faster on the purely physical setup compared to the purely virtual configuration, but there was no significant difference in task accuracy.

Previous research has also investigated the qualitative factors that influence holographic display usage. Medeiros *et al.* investigated the usefulness and configuration of holographic displays in shared transit spaces [44]. Given the space constraints of public transit, it is impossible to configure large physical displays, which motivates the usage of MR. They found placement of holographic displays was driven by social norms and etiquette, followed by comfort. A majority participants noted the benefit of being able to place displays on physical structures within the environment and orient the displays to face them. Ng *et al.* investigated the placement of holographic displays in airplane environments [49]. In their study, participants reported that the ability to create multiple displays of varying sizes in MR increased their ability to multitask compared to singular displays. Participants had varying preferences for configuring displays horizontally or vertically, shedding light on the benefits of MR to provide personalized display setups. In contrast to prior work, our study is not aimed at replacing all traditional displays with MR, but instead seeks to understand how MR can integrate within, and complement existing displays as an augmentation technology in a hybrid information ecosystem.

3 Study

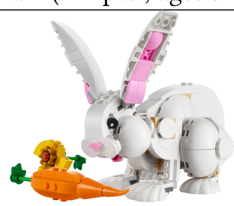
To explore the utility of MR for multi-task work in MDEs, we performed a within-subjects experiment that compared three display modalities: a large display only (*LD*), representing the modality that is most commonly used to provide shared information in these settings, a tablet display combined with a large display (*T+LD*), capturing what is commonly used to provide individual sub-task information, and an MR configuration that allows individual sub-task information to be represented as virtual screens combined with a large display (*MR+LD*). The study's design attempts to bridge a critical gap between prior work that has studied MR for single task guidance and situate it within an environment that is representative of MDEs where users have to perform multiple split-attention tasks in the presence of distractors.

3.1 Experimental Task

Our experimental task required users to assemble two Lego sets per display modality. Table 1 provides descriptions of each set, including number of pieces and age ratings. The Lego sets were placed on a table constructed by combining two modular conference room tables. Each table was 1.5 by 0.6 meters creating a uniform workspace of size 1.5 by 1.2 meters. The Lego sets were placed on opposite ends of the work surface. All pieces were laid out on the work surface (see Figure 1). All tasks would start when the first instruction page was shown on the modality.

We chose a Lego based task on extensive prior work that has used Lego set assembly of similar size and complexity for assembly tasks in lab-based user studies [3, 8, 34, 62]. In particular, Lego building is an ideal task because it does not require domain expertise to guide task decisions but still requires spatial reasoning, visual search, sequencing, attention prioritization, and bi-manual fine-motor manipulation. These are all skills exercised in manual tasks typically performed in MDEs. However, the Lego task did not present any fundamental touch interaction constraints, unlike the previously illustrated cooking example. We chose to not replicate these interaction constraints, as users would have been forced to change their typical behavior for the sake of using the technology (i.e. wash their hands or put down tools during a repair), breaking the authentic context of the task.

Table 1. Lego set names and piece counts for each modality.

 Taxi (124 pcs., ages 7+)	 Unicorn (145 pcs., ages 7+)	 Shark (230 pcs., ages 7+)
 Rabbit (258 pcs., ages 8+)	 Nest (232 pcs., ages 9+)	 Dump Truck (177 pcs., ages 7+)

Requiring the user to perform additional actions before touch interactions would have unfairly biased the MR modality.

Generally, when a task necessitates the use of an MDE, users are constantly performing context switches and having to maintain spatial awareness with distracting content. Distractors in MDEs usually take the same form as distractors in normal desktop environments: notifications from other applications and content from prior tasks on other displays [13, 59]. To replicate this experience, the task had three additional extraneous streams of Lego instructions. To recreate context switches, the task required participants to switch between the two Lego sets at random intervals ranging between 15 seconds and 1 minute. As the focus of our study was to understand the impact of interaction affordances provided by each modality, we controlled for other factors to eliminate confounds (interpersonal dynamics and non-task related communication).

3.2 Participants

Following established protocols for sample size estimation, we conducted a within-factors repeated measures analysis of variance (RM-ANOVA) sample size calculation using G*Power [17]. Our calculation yielded a target sample size of 18, assuming a medium effect size (Cohen's f = 0.25), α =.05, and 80% power. We used a higher correlation value among the repeated measures, ρ = 0.7, as the task performed in each experimental condition had high similarity.

We successfully recruited 18 participants. Recruitment was done through in-class solicitation presentations and snowball sampling. The study was approved by our institution's ethics review board. The inclusion criteria required participants to be over the age of 18, have English fluency, normal vision or vision corrected to normal, and prior experience with Lego assembly. Before beginning the experiment, participants were asked to complete a short demographics questionnaire administered on an iPad tablet (Qualtrics). The questionnaire included questions about gender, age, previous experience with AR and VR on a five-point Likert scale that ranged from "little to none" to "extensive experience," perceived societal value of AR/VR technology on a scale from "little to no value" to "significant value." After completing the questionnaire, participants entered the study space and sat at the work surface in front of a large screen, measuring 98 inches diagonally (Figure 1).

Participants were students or faculty from our institution, a large research university in the United States. 11 were computer science undergraduate students, 2 English undergraduate students, 2 information science doctoral students, 1 chemistry undergraduate student, 1 physics undergraduate student, and 1 computer science lecturer. 10 participants self-identified as male, 7 female, and 1 non-binary. Ages ranged from 18 to 37, with a mean age of 22.28 (σ = 5.79). Self-reported experience with AR ranged from 1 to 3 (out of 5), with a mean of 1.94 (σ = 0.80). Experience with VR ranged from 1 to 4 (out of 5), with a mean of 2.50 (σ = 0.99). Perceived value of MR ranged from 2 to 5 (out of 5), with a mean of 3.94 (σ = 0.80). Participants were not previously familiar with the concept of MDEs and did not have it explained to them before participating. Participants received $15 in compensation.

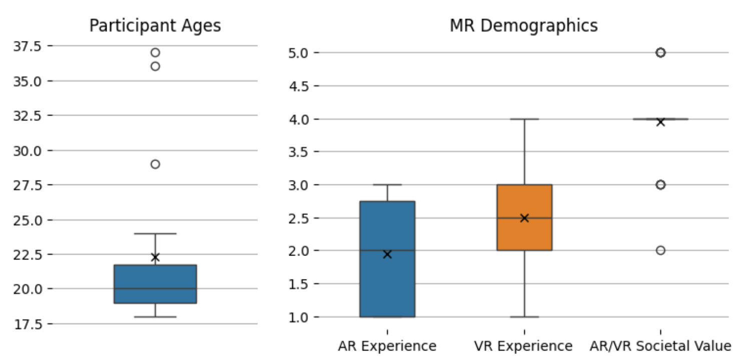


Fig. 2. Participant demographic survey results. Most participants were inexperienced with augmented and virtual reality, but still believed the technology has value for society.

3.3 Procedure

The 3 experimental conditions were performed sequentially with two simultaneous tasks for each modality. To counter for order effects, we counterbalanced the ordering of display modalities using a 3×3 Latin square. Each task was 5 minutes long. Participants were instructed to build as much as possible on both sets within the allotted time. Following each condition, the participant engaged in 5-10 minutes of post task questionnaires. The participant and researcher then engaged in a semi-structured interview which took approximately 10 minutes. The majority of participants completed all study activities within an hour.

Within each modality, participants were first instructed how to use the interface for the experimental tool. After the experimenter demonstrated basic usage of the technology, participants were given time to familiarize themselves with the modality before choosing to start the task. The time each participant chose to take for familiarization varied, but all periods were under 5 minutes. The task began with the participant building as much as possible of the first set. They were told to consult the instructions using the provided modality for the current experimental condition (e.g., *LD*, *T+LD*, or *MR+LD*). At a random interval (between 15 seconds and 1 minute) an audio chime was played which indicated that the participant should shift their attention to focus on building the other set. Participants were instructed about the meaning of the audio chime before beginning the experiment. This process repeated a variable number of times over the course of the five-minute task, with participants switching between the first and second Lego sets. Participants were allowed to reconfigure the modality, if appropriate, to display or enhance the relevant instructions for the current set. We explicitly chose to not control for the position of virtual content, as we believe this would break the authentic use context for MR devices. Specifically, MR provides value by allowing content to be freely positioned and spatially anchored based on a user's preference, context of their specific task, and style of work. This also allowed users to experience the constraints of the technology, such as field of view, better representing its authentic use cases. Similarly, we did not prevent users from repositioning the tablet, as it is also natural for users to move these devices (and the displayed content) in normal use. Interaction control and advancement of the instruction sequences is discussed in subsection 3.5.

3.4 Measurements

The number of instructions completed for both sets within each display condition was automatically recorded in software. This produced two numbers for each modality, one for each Lego set. Instructions were shown to participants from the pages of the Lego instruction booklet, and each page often included multiple brick placement actions. To standardize performance across sets, we counted the number of individual brick placements for each instruction page and summed the total number of placement actions for each set within a display modality. We chose this metric to evaluate performance by following methods of prior work with our experimental design. Specifically, previous studies investigating Lego construction as an experimental task have used "time until model completion" as a performance metric [8, 62]; however, this study did not require participants to finish construction of the models due to switching between two of them. Since the number of Lego bricks in a completed model is fixed, "time until model completion" can be thought of as an analogue for "number of bricks placed per time interval". Instead of maintaining a constant number of bricks that must be placed, our metric fixes the the time interval to 5 minutes. This allows for the most effective comparison to prior work. Questionnaire data included subjective cognitive load (NASA-TLX [22]), System Usability Score (SUS) [2], and Situational Awareness Rating Technique (SART) [63]. The semi-structured interview asked the participant to reflect on their modality rankings, attention distribution, performance per modality, and situations in the their everyday life where a specific modality may be useful.

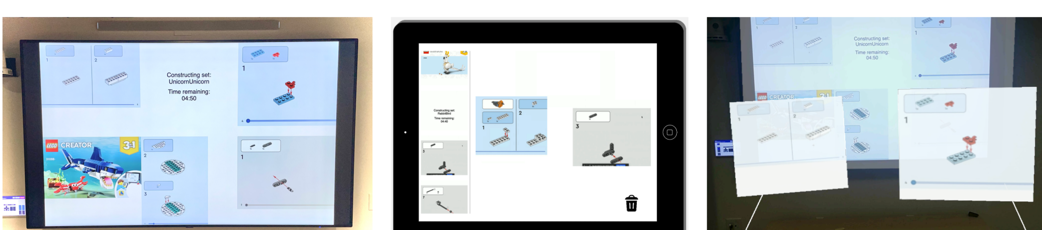


Fig. 3. The three display modalities investigated in the study. The left image shows the large display (*LD*) modality, where content is placed on a shared display with instructions for multiple tasks and users. The middle image shows the tablet interface as part of the *T+LD* modality, where the large display is segmented into individual windows which can be positioned and scaled on the tablet screen. The right is the mixed reality (MR) interface as part of the *MR+LD* modality, where content relevant to the current task is displayed in world-anchored virtual windows.

3.5 Environmental Setup and Tool

No current commercial system or research prototype enables the capture, segmentation, and subdivision of graphical interfaces from common desktop applications into world-anchored virtual representations. To be able to study within the context of MDEs, we developed a new system with these capabilities. We designed and built an experimental tool that allows users to segment a captured display along predefined boundaries, displaying the resulting subdivisions as interactive tiles represented as virtual world-anchored objects or shown on a tablet screen. The experimental tool comprises four parts: the instruction application, the signalling and streaming server, the MR streaming client, and the tablet streaming client. An overview of how information is transformed in the tool can be seen in the video associated with this work and the interfaces for each modality are shown in Figure 3.

3.5.1 Instruction Application (LD Modality). The instruction application was written in HTML, CSS, and JavaScript and designed to be run in a web browser. The application presents 6 information items in a 2×3 grid. This was chosen to mimic a variety of shared display layouts common in MDEs. To support our experimental design, most information items represented instruction sequences for different Lego sets.

The application, by default, operates as the *LD* modality for our experimental design. At launch, the instructions are pre-loaded and a member of the research team selects the two Lego sets to be used for the experimental task. Of the 6 information components, 5 are instruction sequences and 1 shows the current Lego set and the time remaining for the experimental task. 3 sequences are of random Lego sets not related to the participant's task. The unrelated sequences advanced automatically at a randomized rate of one instruction figure every 1-10 seconds. 2 sequences are of the Lego sets assigned to the participant for construction. All segment positions were randomized to one of the six positions for each experimental condition.

To start the task, the experimenter presses the space bar to show the first instruction and initiate the automatic advancement of the unrelated sets. The experimenter advances the currently active Lego set to the next instruction by pressing the space bar. During the experimental task, participants switched between construction of the two assigned Lego sets at a random time interval between 15-60 seconds. The signal to switch was given as an auditory chime and reflected in the task status segment. After time expired, the system displays the number of instruction pages that were completed for each set.

3.5.2 Streaming Server. The streaming server is a web-based application developed in HTML, CSS, and JavaScript. The server streams a capture of the computer's screen using WebRTC, a low-latency streaming protocol. WebRTC requires a signalling component that is shared between peers to exchange information about available communication channels (known as ICE candidates) and a connection string [27]. The signalling component is implemented in NodeJS with the Socket.IO communication protocol. For this study, the screen was captured and streamed at a resolution of 1280 pixels by 720 pixels. Before streaming, the experimenter segments the screen by drawing boxes around a displayed preview or selecting the default segmentation option configured for the experiment setup (6 equally-sized segments). Video is streamed to clients in a single stream; segmentation is performed on the client with bounds provided by the streaming server.

3.5.3 MR Streaming Client. The MR streaming client was built for the Microsoft HoloLens 2. The HoloLens 2 was chosen for several reasons. Firstly, the translucent display of the device allows participants to view their task in a manner more representative of their unobscured vision. Second, the HoloLens 2 is considered to be a state-of-the-art device and is widely used in the literature [31, 33, 50].

We used the WebXR standard with the THREE.js JavaScript library to build the client. A web-based format for the application was chosen so the system could be used on multiple MR headsets during testing without additional configuration and for better WebRTC support. Once initialized, the MR system must first be calibrated by outlining the bounds of the large display. This is accomplished by the experimenter performing a pinch gesture at the four corners of the display where the instruction application is being presented. After calibration and connection to the server, the MR client displays a faint virtual element over the large display as a guide, known as the guide element. Segments of the screen can then be "virtualized" (selected to be replicated as virtual screens) through simple pinch and pull gestures of the instruction sequences on the guide element. This creates a 3D virtual clone of the selected instruction sequence. Only one clone of each display segment can be created at once. If a virtual replicate becomes difficult to manipulate, the user can destroy the previous version through a pinch and pull gesture on the same instruction sequences from the guide element. A virtual replicate of an instruction sequence can be deleted by positioning it near its original location on the guide element. Virtual replicates of instruction sequences update in near real-time ($\leq$ 250ms) with changes of the source display.

3.5.4 Tablet Streaming Client. In some MDEs, users integrate tablets that can subdivide elements of shared information. This affords the ability for users to focus attention on sub-task-related information [51, 56, 61]. To effectively model this affordance, we included a tablet in the *T+LD* modality that allows participants to view a subset of information from the large display.

The tablet streaming client has similar construction to the MR streaming client; it is also a web application, allowing for cross-platform deployment and sharing of software components. After establishing connection with the server, the tablet client builds the left sidebar interface using the provided segmentation bounds. When the user taps one of the preview segments presented on the left sidebar, a clone of the segment is created and placed in the "playspace," the large area in the middle of the tablet screen. Once segment clones have been created in the "playspace," the user may manipulate them with dragging and scaling gestures. A user may have multiple different segments in the "playspace" at one time. However, similar to the MR client, multiple clones of the same segment are not allowed; if the user attempts to clone a segment that is already extant in the "playspace", the older clone is destroyed. When a user wants to remove a segment from the "playspace," they can drag the segment near the trash can at the bottom right corner. Like the MR streaming client, the contents of the sidebar and segments in the "playspace" update in real-time with changes in the source display.

3.5.5 Equipment Used in Study. The large display used in the experiment was a 98" NEC C981Q. The streaming server was hosted on an Apple MacBook Pro (MacOS version 14.0). The MR streaming client was run on a Microsoft HoloLens 2 (OS version 22621.1272) in the Edge Browser (version 121.0.2277.128). The tablet streaming client was run on an Apple 9th Generation iPad (iPadOS version 17.3.1) in the Safari browser. All devices were connected via a local wireless network.

4 Results

All statistical analyses were conducted using SPSS (version 28.0.1.1) along with Python Pandas and SciPy libraries. If the data satisfied the assumptions of normality and sphericity, ANOVA was used. In cases where these assumptions were not met, the Friedman test was used. Note that for all correlation measurements we performed two-tailed tests for significance. We transcribed all interviews using an automated transcription service and made manual corrections as needed to rectify transcription errors.

4.1 Task Performance

The total number of assembly actions completed was calculated by summing the bricks placed across both sets within each task. The *T+LD* modality exhibited the highest performance on average, ($\mu = 22.50,\ \sigma = 8.07$), followed by *LD* ($\mu = 22.44,\ \sigma = 11.62$) and then *MR+LD* ($\mu = 20.28,\ \sigma = 6.94$),

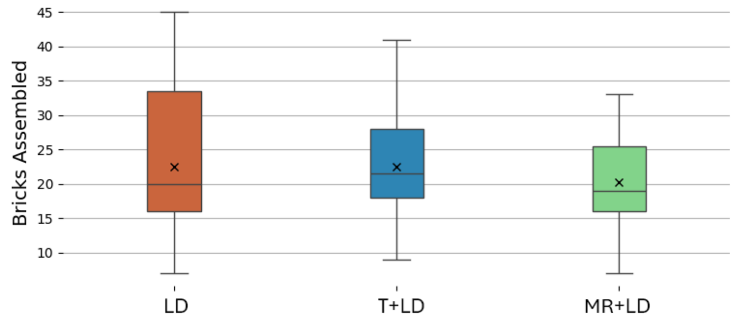


Fig. 4. Performance, measured as number of Lego bricks placed. There was no significant difference between the three modalities.

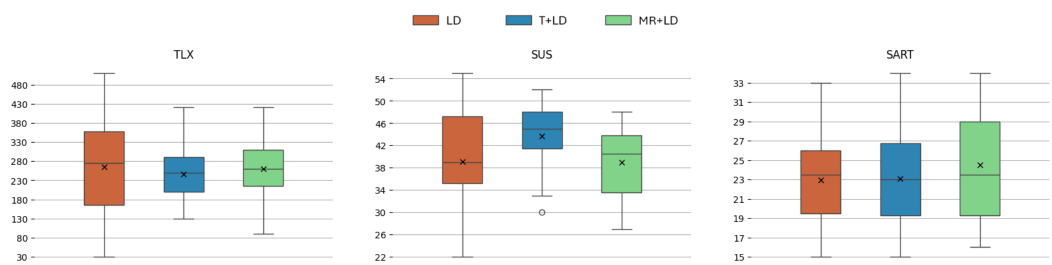


Fig. 5. TLX (cognitive load), SUS (system usability), and SART (situational awareness) scores by modality as a box plot. No significant difference was found between modalities for any of the three measures.

as shown in Figure 4. A one-way repeated measures analysis of variance (RM-ANOVA) test yielded no significant difference in performance between the three modalities.

4.2 Questionnaire Responses

Task load (NASA-TLX) was collected on a scale from 0 to 100 in 10 point increments. The total score was calculated by summing the responses for each question, with the performance question ("How successful were you in accomplishing what you were asked to do?") negated by subtracting its value from 100. A higher total score represents a higher task load. The *T+LD* modality had the lowest score ($\mu = 246.67, \ \sigma = 74.60$), followed by *MR+LD* ($\mu = 260.00, \ \sigma = 89.11$), and *LD* ($\mu = 264.44, \ \sigma = 122.05$).

The System Usability Score (SUS) was computed by summing the responses for each question, with negative usability questions ("I found the system unnecessarily complex", "I found the system very cumbersome to use", "I needed to learn a lot of things before I could get going with this system", "I thought there was too much inconsistency in this system", and "I was distracted by the actions of the system") negated by subtracting their values from 6. A higher score represents a higher assessment of usability. *T+LD* scored highest ($\mu = 43.67, \ \sigma = 6.71$), followed by *LD* ($\mu = 39.06, \ \sigma = 10.04$), and *MR+LD* ($\mu = 38.94, \ \sigma = 5.94$).

The Situational Awareness Rating Technique (SART) scores were summed, with positive questions ("How aroused are you while completing the task?", "How familiar were you with the task?", and "How much information did you gain during the task?") negated by subtracting their values from 8. The lower the SART score, the greater the assessment of situational awareness. *LD* had the lowest score ($\mu = 22.94,\ \sigma = 4.83$), followed by *T+LD* ($\mu = 23.06,\ \sigma = 5.34$), and *MR+LD* ($\mu = 24.50,\ \sigma = 5.58$).

We conducted a one-way repeated measures analysis of variance (RM-ANOVA) on the three sets of questionnaire responses (TLX, SUS, and SART) and found no significant difference between the three modalities. All results are visually presented as box plots in Figure 5.

4.3 Modality Rankings

After completing the tasks for all modalities, participants were asked to rank the modalities according to their preference. Ranking were recorded with 1 being the most preferred and 3 being the least preferred. Visualized in Figure 6, the *T+LD* modality was the highest rated with a mean participant ranking of 1.89 ($\sigma = 0.90$); 8 participants ranked it first, 4 participants ranked it second, and 6 participants ranked it last. The *LD* modality had a mean participant ranking of 2.00 ($\sigma = 0.84$); 6 participants ranked it first, 6 participants ranked it second, and 6 participants ranked it last. The *MR+LD* modality had a mean participant ranking of 2.11 ($\sigma = 0.76$); 4 participants ranked it first, 8 participants ranked it second, and 6 participants ranked it last.

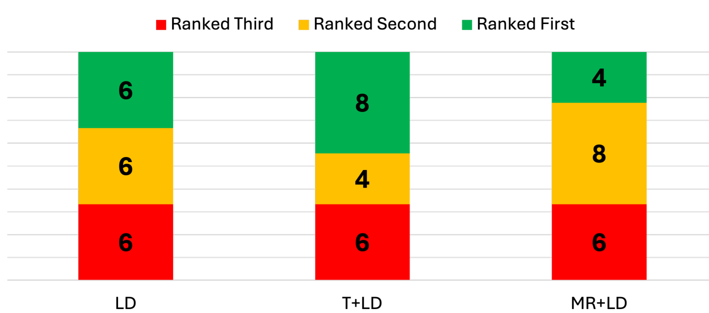


Fig. 6. Participant rankings by modality. Participants ranked the *T+LD* modality first most often, then the *LD* modality, followed by the *MR+LD* modality.

4.4 Performance and Preference

We compared modality preference rankings with each participants' top performing modality. For each modality, we divided participants into two subgroups: the participants who placed the most blocks in that modality and those who did not. We then compared the preference ranking for that modality between the two subgroups. Placing the most bricks in *MR+LD* was linked to preferring *MR+LD* over the other modalities compared to those who placed the most bricks in any other modality ($\mu = 1.33,\ \sigma = 0.58$ vs $\mu = 2.25,\ \sigma = 0.68;\ t(17) = 2.17,\ p < 0.05$). Similarly, placing the most bricks in the *LD* modality was linked to a preferring the *LD* over the other two modalities, compared to those who placed the most bricks in another modality ($\mu = 1.50,\ \sigma = 0.53$ vs $\mu = 2.27,\ \sigma = 0.90;\ t(17) = 2.15,\ p < 0.05$).

4.5 Other Relationships Among Measures

Preferring *LD* was correlated with a lower TLX ($r^2 = 0.144,\ p < 0.01$) and higher SUS ($r^2 = 0.35,\ p < 0.01$) for that modality. Similar statistically significant correlations were not found for the other modalities. Modality preference and SART did not have a significant correlation. Increased situational awareness (lower SART) was correlated with better performance in *LD*

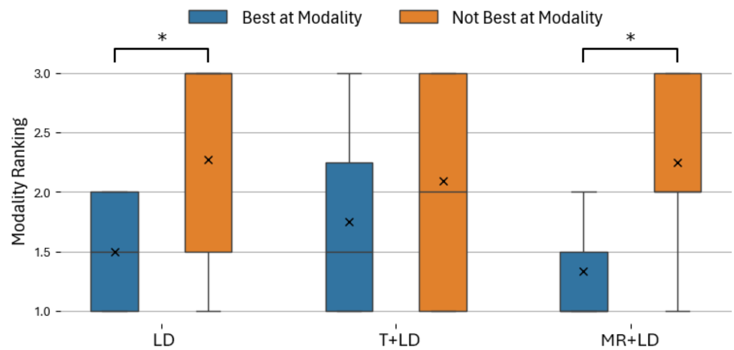


Fig. 7. Participant ranking for the modalities they best performed in. There was a significant difference in ranking for best performers in the *LD* and *MR+LD* modalities, but not *T+LD*.

($r^2 = 0.31$, $p < 0.05$) and *MR+LD* ($r^2 = 0.23$, $p < 0.05$) modalities. Situational awareness and performance on the *T+LD* modality were not significantly correlated.

Participants who performed well in one modality tended to perform well across modalities, based on pairwise comparisons of performance between two modalities (*LD-MR+LD*: $r^2 = 0.45$, $p < 0.01$, *LD-T+LD*: $r^2 = 0.59$, $p < 0.001$, *MR+LD-T+LD*: $r^2 = 0.52$, $p < 0.001$). Participants who spent more time interacting with the tablet tended to place more blocks in the *T+LD* modality ($r^2 = 0.36$, $p < 0.01$).

No other significant correlations were found, including between rank, performance, previous MR experience, and experience metrics (TLX, SUS, and SART).

4.6 Movement of Virtual Elements

We conducted analysis of the telemetry data for both the *MR+LD* and *T+LD* modalities. We found that, in the *MR+LD* modality, participants positioned virtual elements an average of 1.245 meters ($\sigma = 0.574$ meters) apart from one another and moved virtual elements an average total distance of 3.350 meters ($\sigma = 2.780$ meters) after creating the virtual elements. With the *T+LD* modality, participants adjusted virtual elements an average of 206.474 times ($\sigma = 220.696$), and with the *MR+LD* modality, participants adjusted virtual elements an average of 40.278 times ($\sigma = 220.696$).

4.7 Semi-Structured Interview Responses

Based on the flow of the semi-structured interviews, the conversation led in a direction that allowed the interviewer to ask whether participants ranked the modalities based on personal comfort or perceived task efficiency for 15 participants. 10 participants indicated perceived task efficiency, 3 participants indicated personal comfort, and 2 participants indicated both factors equally. All participants were asked which modality allowed them to best focus their attention. 6 indicated *LD*, 6 indicated *T+LD*, 5 indicated *MR+LD*, and 1 indicated a tie between *MR+LD* and *LD*. Of the 6 that chose *MR+LD*, a follow up question was asked to determine if the benefits of MR outweighed the challenges of using the modality. 4 indicated yes, 2 indicated no.

Participants were asked to reflect on their performance using all three modalities. They were given the opportunity to express if they felt outside factors beyond the modality impacted their performance. 7 participants reported that they believed their perceived differences in difficulty between the Lego sets could have contributed to their performance. For instance, participant 8 stated their performance was "*definitely not just purely the [modality]. I hate to bring up the taxi set* *but that felt so much easier than all the other ones... and then like there's probably set [differences], especially in terms of colors and searching for pieces like the ones that were less homogeneous in color...*" 3 participants stated their performance may have improved as they became more comfortable performing the Lego assembly task. Participant 11 reported that "*[with] the third one I did... by that point I was more used to moving or using Legos and finding pieces.*".

Interviews asked participants to describe the benefits and drawbacks of each modality. Transcriptions of these responses were analyzed through a general thematic analysis [9] process. Recordings totaled 2 hours, 56 minutes, and 46 seconds. They were transcribed using the Whisper Transcription application[3] using the "Medium English" AI speech recognition model. The primary author reviewed the transcripts and synthesized themes from the dialog. Themes were shared and revised by the other authors on a subset of the transcriptions. The primary author used the feedback to revise the transcript coding.

For the *LD* modality, participants indicated its comprehensiveness (5 participants), minimized interaction (5), ease of use and perceived comfort (5), clarity (3), and capacity for collaboration (3) as benefits. For instance, Participant 3 stated *"I think it's just because like there's only one place to look [for the large display] and I think when I have the tablet and the MR there was the screen and there was the other screen and it was like the two Lego sets so I think it's just more straightforward."*.

Participants indicated the large display's physical size (9 participants), lack of personalization and poor information layout (5), increased cognitive load (4), and limited field of view (3) as drawbacks. For instance, Participant 9 expressed concern about large displays being in all work environments: *"...the large display might be in an area in an environment where that's not possible to have. There could not be, like, a wall or a power source to display that. So that might be better if you have [MR] instead."*

Participants indicated the *T+LD* modality's interactivity and personalization (11 participants), proximity to the tasks (7), familiarity (6), and portability (4) as benefits. For example, Participant 14 stated *"I feel like the the capability that it had of having multiple tabs open on the same screen would be useful if not everyone needs to be looking at the same information."* Specific drawbacks of the *T+LD* modality mentioned by participants included interaction issues or hand occupation (11), its limited view capability (5), and its physical space requirements (2). Participant 15 stated *"I had to keep, like, readjusting the screen and, like, zooming in... and then click on the next one and then I'd have to come back and forth between the two, like, tasks."*

Mentioned benefits of the *MR+LD* modality included its limited interaction and hands-free nature (14 participants), its personalization and increased task attention (12), and its information detail, including depth capability (6). Participant 16 noted *"...the headset has the advantage that you have your two hands free and you don't have anything [in] the way."* Noted drawbacks of the *MR+LD* modality included system concerns (like battery life and cost) and physical discomfort (9), image issues like color saturation (8), lack of familiarity and interaction issues (8), and obstruction of real-world view and lack of outside attention due to immersion (7). For instance, Participant 8 stated *"I feel like the mixed [reality] would have been my number one if you'd given me like 20 more minutes in it."*

5 Discussion

This work provides insights into the state of research on integrating MR into MDEs. Below, we situate how our work extends the understanding beyond the literature on MR in information-rich environments. Our study particularly draws attention to the relationship between preference and performance for MR-guided tasks and the importance of natural interaction for success with MR in

[3]https://apps.apple.com/us/app/whisper-transcription/id1668083311

these environments. These insights are useful to guide the continued development and design of MR-based tools and applications for these ecosystems and beyond.

5.1 MR as a Display Modality

5.1.1 Prior Research. The use of MR in the context of task-guided work has been well studied in prior research, with inconsistent lessons on the impact of the modality. Specifically, objective measures of performance, task load and situational awareness are variable in significant difference. Tang *et al.* found no significant difference in task performance when instructions were presented in MR compared to being presented on a large monitor. However, their study did find subjective task load to be lower with MR for a Lego assembly task [62]. In contrast, Blattgerste *et al.* also had participants perform a Lego assembly task and found that they performed worse in MR compared to paper instructions, despite making fewer errors [8]. Even though these two works used relatively similar methodologies, they do not come to a shared conclusion on the effects of MR on task performance. Our work provides a similar result for task performance to Tang *et al.*, although we did not find a similar decrease in cognitive load. These inconsistencies seem to permeate the broader field. Daling and Schlittmeier found some contradictory results in their detailed review of MR task-guided work [10]. They attribute these contradictions to diversity of the hardware used and types of the manual assembly tasks.

5.1.2 Insights Gained. Toward understanding **RQ1**, the study did not find significant performance differences between modalities. This, at least within the physical tasks studied, showed the use of MR displays in place of traditional displays does not have a performance impact. This establishes an important early benchmark toward the integration of MR display modalities in future MDEs. Users with little training were able to utilize an MR information display as effectively as traditional display surfaces for information gathering and integration while context switching. We also found that cognitive load and situational awareness were not affected using an MR display. This demonstrates that the integration of MR into MDEs is mechanically easy but is not sufficient for performance advantages of MR. When combined with results of prior research, the value of MR's value in MDEs should better explore and measure the specific benefits of 3D content that MR can provide to these hybrid spaces.

5.2 MR's Role in Future MDEs

As previously discussed, there are a multitude of everyday tasks that can benefit from the affordances of MDEs. Yet, MDEs are more difficult to leverage in many interaction-limited tasks. For example, tasks that require gloves or frequent hand cleaning are not conducive to traditional input paradigms like touchscreens. A person preparing a family meal may want to view their recipe, a kitchen timer, an instructional video, and a unit converter simultaneously, but their hands are busy and covered in food. This is one possible reason for the increasing presence of voice-based virtual assistant devices in the kitchen, as they do not require manual interaction [68]. However, this entirely eliminates the visual aspect of presentation and forces the user to *request* information as opposed to having it presented to them. The interaction constraints of the task and the environment make the implementation of an MDE in these spaces difficult.

MR provides the benefits of both voice-based virtual assistants through speech recognition and planar display presentation without the necessity for physical interaction, making it a prime candidate for enabling MDEs in interaction constained environments. Toward understanding **RQ2**, insights from our study suggest that its current form factor may be preventing MR from entering these spaces. To gain access to the core benefits of MR technology, like world-anchored virtual objects and eye tracking, users must boot a separate headset device, put it on, and then configure the device for their intended use case. This is in contrast to most ubiquitous computing devices like smartphones and tablets which can remain stashed away in a pocket or drawer, while being only seconds away from integrating into a user's workflow. Recent developments in MR, like Meta's *Orion* project, demonstrate a generational paradigm shift, not only in the physical form factor of the technology, but the manner in which users traverse the physical-virtual boundary. For everyday use cases, MR headsets are transitioning from immersive experience devices to a tool akin to reading glasses – users can put them on at a moment's notice and wear them for extended periods of time. Additionally, with the advent of accessible, affordable spatial computing technology, environments that were previously only "multi-display" can become "multi-information", where different modes of information display (3D virtual elements, haptic feedback sources, large displays, touch devices, etc.) are integrated together to provide the best experience for each individual task or data point. Such an environment would allow information to exist as a first-class citizen and break away from the boundaries of its modality.

What remains unclear is the technique by which these modalities should be integrated to create a unified multi-information environment. Some commercial offerings like Meta Horizons Workrooms[4] attempt to replace traditional displays by entirely virtualizing a worker's environment. Alternatively, like volumetric windows on the Apple Vision Pro[5], they create new forms of interactivity. Other MR displays, including the one studied, have taken inspiration from planar displays and traditional window systems. With our research prototype system, we attempted to leverage users' prior familiarity with gestures like "click-and-drag" to provide a smooth path for them to convert physical displays to virtual displays. The form factor of MR will dictate the method in which users transition information across modalities. If MR technology continues to require users to follow an intentioned startup and shutdown routine, the interaction will likely remain isolated and difficult to integrate into a larger workflow. For instance, there are limits as to how quickly a user can initialize and configure modern headset devices, especially if they require users to connect an external battery pack, plug the device into a computer, or calibrate certain display or interaction components. This is in contrast to a light-weight form factor akin to reading glasses, which users could put on and take off as they need assistance. With this lighter-weight form factor, perhaps interactions could become more episodic and integrated. The capacity for MR to integrate with other modalities is undoubtedly dependent on the environment, task, and users. Work that limits or restricts touch interaction, like sterile surgery, may promote faster adoption of MR due to less existing cross-modality interaction. This may shed light on the recent success of MR research within surgery [5, 29, 30, 40]. As our results show, there is opportunity for additional innovation in cross-modality interactions, and this remains an open question for future work. The integration of MR into collaborative MDEs is even more complex, as multiple virtual elements in a shared space can obscure one another [32]. Overall, the transition from traditional planar information presentation systems to novel mixed reality displays will happen, but it will require much more advanced and perhaps yet invented paradigms for manipulating information within, between, and distributed throughout display modalities that comprise MDEs.

5.3 Performance, Preference, and Interaction

When we sorted users by their top performing modality, we found an interesting confound. Top performers in the *LD* and *MR+LD* modalities were more likely to prefer the modality in which they performed best. In interviews, most participants indicated their preference rankings were based on perceived task efficiency despite not being told any objective feedback on their performance.

[4]https://forwork.meta.com/horizon-workrooms/

[5]https://developer.apple.com/documentation/visionos/creating-a-volumetric-window-in-visionos

It is possible that a participant's higher performance in a modality may have resulted in a more enjoyable experience performing the Lego assembly task. After all, the main purpose of Lego sets is enjoyment.

The study observation may also be highlighting that the fundamental tension between preference, performance, and environmental constraints are likely balanced differently across varied approaches and practices. For instance, an insight from our study suggests that users who prefer to position task information directly within the their field of view valued the affordances of MR. Integration of MR in MDEs may allow this affordance to be practical, as it can be provided without impeding the ability of coworkers to perform work in a shared workspace (e.g., shared work around a patient on an operating table). Future studies are needed to explore the benefits of MR in concurrent tasks with a combination of shared and personalized information.

The positive relationship between increased tablet manipulation time and performance may be indicative of the interaction benefits offered by the modality. In the study, participants did not have to interact with the tablet to advance the task, as the pace of instructions was controlled via verbal instruction. We expected that most participants would have simply positioned the two segments showing instructions for the active Lego sets at the beginning of the experiment and cease interaction. However, performant participants chose to take an active role in managing the limited display resources available. Interactions among high-performing participants usually took the form of repositioning, zooming, deleting, and recreating segments. The link between increased interaction and performance suggest the need for more usable and natural interaction techniques for managing planar content is needed for MR. The techniques used in the study were adapted from a common toolkit[6]. Repositioning content required hard to remember and temperamental gestures that are objectively more complex compared to the direct manipulation interaction studied on the tablet. When more natural, effective interaction techniques become available, a repeated comparison study should continue to investigate the link between interaction time and performance for the *MR+LD* modality.

Further addressing **RQ1** and **RQ2**, preference may have been driven by the specific task affordances of each modality. For instance, Participant 6 stated *"I liked the tablet the most because... I could control it the most..."* and mentioned the presence of familiar signifiers: *"It also had the trash can, which I understood and it was just easy. It was the easiest to work with"*. Participant 15 enjoyed the world-anchored virtual displays of MR more than the traditional screens *"because I have to look up into the screens and, like, look away from the task I'm doing rather than just [continuing to look] at the task while also having the screen right next to me"*. Some participants enjoyed the static nature of the large display: *"it didn't move around... I would just look at where I knew the information was and then look down. It felt easier to access the information." - P18*. This further demonstrates the task dependent nature of the integration of MR into MDEs. Although the affordances mentioned by participants were specific to the task, it shows that the abilities of each modality were valued differently between users.

6 Limitations

We acknowledge several limitations in our study. First, almost all participants in this study were college-aged students in the same geographic environment. Future investigations should likely expand their user base to include both MR-inexperienced and MR-experienced users, as well as users with varying MDE experience. This will likely become more achievable as the general population begins to engage with MR using more accessible or available hardware such as the Apple Vision Pro. Participants were unable to rearrange the components of the *LD* modality. As described in

[6]https://threejs.org/

Section 3.1, the task that participants completed was contrived, and abbreviated. We made this decision so we could more easily control the circumstances around the task. A co-located task makes controlling confounding factors such as participant interaction and use of shared physical space more difficult. Extending the results to a co-located context remains future work.

It is also important to note that the task in this study did not utilize the primary benefits provided by MR. Namely, the content shown in MR was not 3D. This was an explicit decision as we wanted to create a task and environment where no modalities were specifically disadvantaged. The addition of 3D content introduces an additional variable that would have likely confounded the results. A comparison study that examines the unique affordances provided by each modality is still needed.

We acknowledge that not all environments or tasks are conducive to MDE use, and not all tasks with pre-existing MDEs are suited for MR integration. For instance, tasks that involve interaction between humans, like communication with customers or communication between a doctor and a patient, may experience detriment if one party uses MR and the other does not. This remains an open question for future work.

There are many usability limitations with the HoloLens 2. According to prior work, ideal MR systems should have a resolution of 200 megapixels, a full field of view of 165 by 175 degrees, and a mass in the 10s of grams [25]. The Microsoft HoloLens 2, however, has a 4.4 megapixel display, a diagonal field of view of 52 degrees, and a mass of 566 grams. An optimized MR experience may have yielded different results.

7 Conclusion

This work investigates the effects of different display modalities (MR, tablets, and large displays) on user preference and task performance within MDEs. Our results build on previous work investigating the integration of MR into these spaces. We shed light on the complexity of augmenting traditional displays with MR technologies. Preference, interactivity, and the nature of the task are first-order considerations in the architecture of MDEs. Our findings inform design decisions for maximizing efficiency and comfort within these environments, guiding the development of MR-based tools.

Acknowledgments

This work has been supported by NSF Grant #2139321 and a generous private gift from the Kolvachick Family.

References

[1] Sriram Karthik Badam, Andreas Mathisen, Roman Rädle, Clemens N. Klokmose, and Niklas Elmqvist. 2019. Vistrates: A Component Model for Ubiquitous Analytics. *IEEE Transactions on Visualization and Computer Graphics* 25, 1 (2019), 586–596. doi:10.1109/TVCG.2018.2865144

[2] Aaron Bangor, Philip T Kortum, and James T Miller. 2008. An empirical evaluation of the system usability scale. *Intl. Journal of Human–Computer Interaction* 24, 6 (2008), 574–594. doi:10.1080/10447310802205776

[3] João Barroso, Luís Fonseca, Bernardo Marques, Paulo Dias, and B Sousa Santos. 2020. Remote Collaboration using Mixed Reality: Exploring a shared model approach through different interaction methods. In *European Virtual Conference on Computer-Supported Cooperative Work*. 1–4. doi:10.18420/ecscw2020_p01

[4] Andrea Batch, Peter W. S. Butcher, Panagiotis D. Ritsos, and Niklas Elmqvist. 2024. Wizualization: A "Hard Magic" Visualization System for Immersive and Ubiquitous Analytics. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2024), 507–517. doi:10.1109/TVCG.2023.3326580

[5] Assaf Berger, Osamah J Choudhry, and Douglas Kondziolka. 2023. Augmented Reality–Assisted Percutaneous Rhizotomy for Trigeminal Neuralgia. *Operative Neurosurgery* 24, 6 (2023), 665–669. doi:10.1227/ons.0000000000000661

[6] Jacob T. Biehl, William T. Baker, Brian P. Bailey, Desney S. Tan, Kori M. Inkpen, and Mary Czerwinski. 2008. Impromptu: a new interaction framework for supporting collaboration in multiple display environments and its field evaluation for co-located software development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*

(Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 939–948. doi:10.1145/1357054.1357200

[7] Noemi Bitterman. 2006. Technologies and solutions for data display in the operating room. *Journal of clinical monitoring and computing* 20, 3 (2006), 165–173. doi:10.1007/s10877-006-9017-0

[8] Jonas Blattgerste, Benjamin Strenge, Patrick Renner, Thies Pfeiffer, and Kai Essig. 2017. Comparing conventional and augmented reality instructions for manual assembly tasks. In *Proceedings of the 10th international conference on pervasive technologies related to assistive environments*. 75–82. doi:10.1145/3056540.3056547

[9] Virginia Braun and Victoria Clarke. 2012. *Thematic analysis.* American Psychological Association. doi:10.1037/13620-004

[10] Lea M Daling and Sabine J Schlittmeier. 2024. Effects of augmented reality-, virtual reality-, and mixed reality–based training on objective performance measures and subjective evaluations in manual assembly tasks: a scoping review. *Human Factors* 66, 2 (2024), 589–626. doi:10.1177/00187208221105135

[11] Fred D Davis et al. 1989. Technology acceptance model: TAM. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption* 205, 219 (1989), 5.

[12] Hitesh Dhiman and Carsten Röcker. 2024. Does the Medium Matter? A Comparison of Augmented Reality Media in Instructing Novices to Perform Complex, Skill-Based Manual Tasks. *Proc. ACM Hum.-Comput. Interact.* 8, EICS, Article 249 (June 2024), 28 pages. doi:10.1145/3660249

[13] Jakub Dostal, Per Ola Kristensson, and Aaron Quigley. 2013. Subtle gaze-dependent techniques for visualising display changes in multi-display environments. In *Proceedings of the 2013 International Conference on Intelligent User Interfaces* (Santa Monica, California, USA) *(IUI '13)*. Association for Computing Machinery, New York, NY, USA, 137–148. doi:10.1145/2449396.2449416

[14] Christopher Dromey and Erin Shim. 2008. The effects of divided attention on speech motor, verbal fluency, and manual task performance. (2008). doi:10.1044/1092-4388(2008/06-0221)

[15] P Dunston, Xiangyu Wang, Mark Billinghurst, Ben Hampson, et al. 2003. Mixed reality benefits for design perception. *Nist Special Publication SP* (2003), 191–196.

[16] Niklas Elmqvist and Pourang Irani. 2013. Ubiquitous Analytics: Interacting with Big Data Anywhere, Anytime. *Computer* 46, 4 (2013), 86–89. doi:10.1109/MC.2013.147

[17] Franz Faul, Edgar Erdfelder, Albert-Georg Lang, and Axel Buchner. 2007. G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods* 39, 2 (2007), 175–191. doi:10.3758/BF03193146

[18] Alessandro Febretti, Arthur Nishimoto, Victor Mateevitsi, Luc Renambot, Andrew Johnson, and Jason Leigh. 2014. Omegalib: A multi-view application framework for hybrid reality display environments. In *2014 IEEE Virtual Reality (VR)*. IEEE, 9–14. doi:10.1109/VR.2014.6802043

[19] Fernando Garcia-Sanjuan, Javier Jaen, and Vicente Nacher. 2016. Toward a General Conceptualization of Multi-Display Environments. *Frontiers in ICT* Volume 3 - 2016 (2016). doi:10.3389/fict.2016.00020

[20] Tony Gjerlufsen, Clemens Nylandsted Klokmose, James Eagan, Clément Pillias, and Michel Beaudouin-Lafon. 2011. Shared substance: developing flexible multi-surface applications. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 3383–3392. doi:10.1145/1978942.1979446

[21] Ramesh Guntha, Balaji Hariharan, and P Venkat Rangan. 2017. WWoW: World Without Walls Immersive Mixed Reality with Virtual Co-location, Natural Interactions, and Remote Collaboration. In *Intelligent Technologies for Interactive Entertainment: 8th International Conference, INTETAIN 2016, Utrecht, The Netherlands, June 28–30, 2016, Revised Selected Papers*. Springer, 209–219. doi:10.1007/978-3-319-49616-0_20

[22] Sandra G Hart and Lowell E Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183. doi:10.1016/S0166-4115(08)62386-9

[23] Tom Horak, Andreas Mathisen, Clemens N. Klokmose, Raimund Dachselt, and Niklas Elmqvist. 2019. Vistribute: Distributing Interactive Visualizations in Dynamic Multi-Device Setups. 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, 1–13. doi:10.1145/3290605.3300846

[24] Dugald Hutchings, Mary Czerwinski, Brian Meyers, and John Stasko. 2004. Exploring the use and affordances of multiple display environments. *Presented at the Workshop on Ubiquitous Display Environments at UbiComp* (01 2004).

[25] Muhammad Huzaifa, Rishi Desai, Samuel Grayson, Xutao Jiang, Ying Jing, Jae Lee, Fang Lu, Yihan Pang, Joseph Ravichandran, Finn Sinclair, et al. 2021. ILLIXR: Enabling end-to-end extended reality research. In *2021 IEEE International Symposium on Workload Characterization (IISWC)*. IEEE, 24–38. doi:10.1109/IISWC53511.2021.00014

[26] Shahram Izadi, Harry Brignull, Tom Rodden, Yvonne Rogers, and Mia Underwood. 2003. Dynamo: a public interactive surface supporting the cooperative sharing and exchange of media. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*. 159–168. doi:10.1145/964696.964714

[27] Cullen Jennings, Florent Castelli, Henrik Boström, and Jan-Ivar Bruaroey. 2023. *WebRTC: Real-Time Communication in Browsers*. Technical Report. W3C. https://www.w3.org/TR/2023/REC-webrtc-20230306/.

[28] B. Johanson, A. Fox, and T. Winograd. 2002. The Interactive Workspaces project: experiences with ubiquitous computing rooms. *IEEE Pervasive Computing* 1, 2 (2002), 67–74. doi:10.1109/MPRV.2002.1012339

[29] Nicolás Kass, Lucille Cheng, Zhazira Irgebay, Elizabeth Moroni, Lucas Dvoracek, and Jesse Goldstein. 2023. Mixed Reality in the Operating Room: Utility within Facial Feminization Surgery. *Plastic and Reconstructive Surgery–Global Open* 11, 10S (2023), 120. doi:10.1097/01.GOX.0000992516.43935.ee

[30] Talha Khan, Edward G Andrews, and Jacob T Biehl. 2024. Future of Surgical Mixed Reality: Cutting-Edge or Cutting Too Deep?. In *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 203–207. doi:10.1109/VRW62533.2024.00041

[31] Talha Khan, Edward G Andrews, Paul A Gardner, Arka N Mallela, Jeffrey R Head, Joseph C Maroon, Georgios A Zenonos, Dmitriy Babichenko, and Jacob T Biehl. 2022. AR in the OR: exploring use of augmented reality to support endoscopic surgery. In *Proceedings of the 2022 ACM International Conference on Interactive Media Experiences*. 267–270. doi:10.1145/3505284.3532970

[32] Talha Khan and David Lindlbauer. 2024. Don't Block My Stuff: Fostering Personal Object Awareness in Multi-user Mixed Reality Environments. *Proc. ACM Hum.-Comput. Interact.* 8, ISS, Article 526 (Oct. 2024), 24 pages. doi:10.1145/3698126

[33] Talha Khan, Toby S Zhu, Thomas Downes, Lucille Cheng, Nicolás M Kass, Edward G Andrews, and Jacob T Biehl. 2023. Understanding effects of visual feedback delay in ar on fine motor surgical tasks. *IEEE Transactions on Visualization and Computer Graphics* (2023). doi:10.1109/TVCG.2023.3320214

[34] Bui Minh Khuong, Kiyoshi Kiyokawa, Andrew Miller, Joseph J. La Viola, Tomohiro Mashita, and Haruo Takemura. 2014. The effectiveness of an AR-based context-aware assembly support system in object assembly. In *2014 IEEE Virtual Reality (VR)*. 57–62. doi:10.1109/VR.2014.6802051

[35] Clemens N Klokmose, James R Eagan, Siemen Baader, Wendy Mackay, and Michel Beaudouin-Lafon. 2015. Webstrates: shareable dynamic media. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*. 280–290. doi:10.1145/2807442.2807446

[36] Rotem Kronenberg, Tsvi Kuflik, and Ilan Shimshoni. 2022. Improving Office Workers' Workspace Using a Self-adjusting Computer Screen. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 12, 3 (2022), 1–32. doi:10.1145/3545993

[37] Ricardo Langner, Marc Satkowski, Wolfgang Büschel, and Raimund Dachselt. 2021. Marvis: Combining mobile devices and augmented reality for visual data analysis. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–17. doi:10.1145/3411764.3445593

[38] Morgan Le Chénéchal, Thierry Duval, Valérie Gouranton, Jérôme Royan, and Bruno Arnaldi. 2016. Vishnu: virtual immersive support for helping users an interaction paradigm for collaborative remote guiding in mixed reality. In *2016 IEEE Third VR International Workshop on Collaborative Virtual Environments (3DCVE)*. IEEE, 9–12. doi:10.1109/3DCVE.2016.7563559

[39] Gun A Lee, Theophilus Teo, Seungwon Kim, and Mark Billinghurst. 2018. A User Study on MR remote collaboration using Live 360 Video. In *2018 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 153–164. doi:10.1109/ISMAR.2018.00051

[40] Guan Li, Jie Dong, Jinbao Wang, Dongbing Cao, Xin Zhang, Zhiqiang Cao, and Guangming Lu. 2020. The clinical application value of mixed-reality-assisted surgical navigation for laparoscopic nephrectomy. *Cancer Medicine* 9, 15 (2020), 5480–5489. doi:10.1002/cam4.3189

[41] Tin-Hui Lin, Chao-Hsiang Liu, Meng-Han Tsai, and Shih-Chung Kang. 2015. Using Augmented Reality in a Multiscreen Environment for Construction Discussion. *Journal of Computing in Civil Engineering* 29, 6 (2015), 04014088. doi:10.1061/(ASCE)CP.1943-5487.0000420

[42] Antonio Maffei, Michela Dalle Mura, Fabio Marco Monetti, and Eleonora Boffa. 2023. Dynamic Mixed Reality Assembly Guidance Using Optical Recognition Methods. *Applied Sciences* 13, 3 (2023), 1760. doi:10.3390/app13031760

[43] Paul McKay. 2009. *Design of collaborative systems for modern cockpits*. Ph. D. Dissertation. University of Waterloo Waterloo.

[44] Daniel Medeiros, Mark McGill, Alexander Ng, Robert McDermid, Nadia Pantidi, Julie Williamson, and Stephen Brewster. 2022. From shielding to avoidance: Passenger augmented reality and the layout of virtual displays for productivity in shared transit. *IEEE Transactions on Visualization and Computer Graphics* 28, 11 (2022), 3640–3650. doi:10.1109/TVCG.2022.3203002

[45] Sarah Mennicken. 2009. *PersonalChef: Cooking Guidance Through an Interactive Multi-Display Kitchen Counter*. Ph. D. Dissertation. RWTH Aachen University.

[46] Adolfo Munoz, Ana Martí, Xavier Mahiques, Luis Gracia, J Ernesto Solanes, and Josep Tornero. 2020. Camera 3D positioning mixed reality-based interface to improve worker safety, ergonomics and productivity. *CIRP Journal of Manufacturing Science and Technology* 28 (2020), 24–37. doi:10.1016/j.cirpj.2020.01.004

[47] Miguel A. Nacenta, Satoshi Sakurai, Tokuo Yamaguchi, Yohei Miki, Yuichi Itoh, Yoshifumi Kitamura, Sriram Subramanian, and Carl Gutwin. 2007. E-conic: a perspective-aware interface for multi-display environments. In *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology* (Newport, Rhode Island, USA) *(UIST '07)*. Association for Computing Machinery, New York, NY, USA, 279–288. doi:10.1145/1294211.1294260

[48] Miguel A. Nacenta, Samer Sallam, Bernard Champoux, Sriram Subramanian, and Carl Gutwin. 2006. Perspective cursor: perspective-based interaction for multi-display environments. 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, 289–298. doi:10.1145/1124772.1124817

[49] Alexander Ng, Daniel Medeiros, Mark McGill, Julie Williamson, and Stephen Brewster. 2021. The passenger experience of mixed reality virtual display layouts in airplane environments. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 265–274. doi:10.1109/ISMAR52148.2021.00042

[50] Leonardo Pavanatto, Chris North, Doug A Bowman, Carmen Badea, and Richard Stoakley. 2021. Do we still need physical monitors? an evaluation of the usability of ar virtual monitors for productivity work. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*. IEEE, 759–767. doi:10.1109/VR50410.2021.00103

[51] Johannes Pfeffer, Markus Graube, Patrick Reipschlaeger, Stephan Arndt, Leon Urbas, Raimund Dachselt, and Ralph Stelzer. 2015. Towards collaborative plant control using a distributed information and interaction space. In *2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA)*. IEEE, 1–4. doi:10.1109/ETFA.2015.7301587

[52] Wim Pouw, Gertjan Rop, Bjorn De Koning, and Fred Paas. 2019. The cognitive basis for the split-attention effect. *Journal of Experimental Psychology: General* 148, 11 (2019), 2058. doi:10.1037/xge0000578

[53] Dheva Raja, Doug Bowman, John Lucas, Chris North, and Virginia Tech. 2004. Exploring the benefits of immersion in abstract information visualization. *8th Int'l Immersive Projection Technology Workshop (IPT '04)* (01 2004).

[54] Manuel Rebol, Colton Hood, Claudia Ranniger, Adam Rutenberg, Neal Sikka, Erin Maria Horan, Christian Gütl, and Krzysztof Pietroszek. 2021. Remote assistance with mixed reality for procedural tasks. In *2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*. IEEE, 653–654. doi:10.1109/VRW52623.2021.00209

[55] Arvin Christopher C Reyes, Neil Patrick A Del Gallego, and Jordan Aiko P Deja. 2020. Mixed reality guidance system for motherboard assembly using tangible augmented reality. In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations*. 1–6. doi:10.1145/3385378.3385379

[56] Laurel C Sadler, Somiya Metu, and US Army Research Laboratory Adelphi United States. 2017. *Intelligent Command and Control Demonstration Setup and Presentation Instructions*. US Army Research Laboratory. https://apps.dtic.mil/sti/citations/AD1043277

[57] Kevin L Schultz, David M Batten, and Thomas J Sluchak. 1998. Optimal viewing angle for touch-screen displays: Is there such a thing? *International journal of industrial ergonomics* 22, 4-5 (1998), 343–350. doi:10.1016/S0169-8141(97)00087-5

[58] Thomas Seifried, Michael Haller, Stacey D. Scott, Florian Perteneder, Christian Rendl, Daisuke Sakamoto, and Masahiko Inami. 2009. CRISTAL: a collaborative home media and device controller based on a multi-touch display. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces* (Banff, Alberta, Canada) *(ITS '09)*. Association for Computing Machinery, New York, NY, USA, 33–40. doi:10.1145/1731903.1731911

[59] Ben Shneiderman and Benjamin B. Bederson. 2005. Maintaining concentration to achieve task completion. In *Proceedings of the 2005 Conference on Designing for User EXperience* (San Francisco, California, USA) *(DUX '05)*. AIGA: American Institute of Graphic Arts, New York, NY, USA, 9–es.

[60] Paul Stager and Gerald J Laabs. 1977. The effect of divided attention on probe reaction time in multiple-task performance. *Canadian Journal of Psychology/Revue canadienne de psychologie* 31, 4 (1977), 174. doi:10.1037/h0081660

[61] Ignus Swart, Jacobus Venter, Louise Leenen, T Meyer, and Willem Roux. 2012. *Command And Control*.

[62] Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 73–80. doi:10.1145/642611.642626

[63] Richard M Taylor. 2017. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness*. Routledge, 111–128. doi:10.4324/9781315087924-8

[64] Balasaravanan Thoravi Kumaravel, Fraser Anderson, George Fitzmaurice, Bjoern Hartmann, and Tovi Grossman. 2019. Loki: Facilitating remote instruction of physical tasks using bi-directional mixed-reality telepresence. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. 161–174. doi:10.1145/3332165.3347872

[65] 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*. 182–195. doi:10.1145/3379337.3415827

[66] B Turgut. 2018. Ocular ergonomics for the computer vision syndrome. *Journal of Eye and Vision* 1, 1-2 (2018), 1–9.

[67] Viswanath Venkatesh and Fred D Davis. 2000. A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science* 46, 2 (2000), 186–204. doi:10.1287/mnsc.46.2.186.11926

[68] Alexandra Vtyurina and Adam Fourney. 2018. Exploring the role of conversational cues in guided task support with virtual assistants. In *Proceedings of the 2018 CHI conference on human factors in computing systems*. 1–7. doi:10.1145/3173574.3173782

[69] James R Wallace, Stacey D Scott, Eugene Lai, and Deon Jajalla. 2011. Investigating the role of a large, shared display in multi-display environments. *Computer Supported Cooperative Work (CSCW)* 20, 6 (2011), 529–561. doi:10.1007/s10606-011-9149-8

[70] Peng Wang, Xiaoliang Bai, Mark Billinghurst, Shusheng Zhang, Sili Wei, Guangyao Xu, Weiping He, Xiangyu Zhang, and Jie Zhang. 2021. 3DGAM: using 3D gesture and CAD models for training on mixed reality remote collaboration. *Multimedia Tools and Applications* 80 (2021), 31059–31084. doi:10.1007/s11042-020-09731-7

[71] Peng Wang, Xiaoliang Bai, Mark Billinghurst, Shusheng Zhang, Xiangyu Zhang, Shuxia Wang, Weiping He, Yuxiang Yan, and Hongyu Ji. 2021. AR/MR remote collaboration on physical tasks: a review. *Robotics and Computer-Integrated Manufacturing* 72 (2021), 102071. doi:10.1016/j.rcim.2020.102071

[72] Xiangyu Wang and Phillip S Dunston. 2005. System evaluation of a mixed reality-based collaborative prototype for mechanical design review collaboration. In *Computing in Civil Engineering (2005)*. 1–9. doi:10.1061/40794(179)11

[73] M. Weiser. 1993. Ubiquitous Computing. *Computer* 26, 10 (Oct. 1993), 71–72. doi:10.1109/2.237456

[74] Chad Weiss, Mary Freiman, Sylvain Bruni, Danielle Ward, Spencer Lynch, and Kristy Kay. 2021. Information Ergonomics: Expediting Maintenance Workflow Using Mixed Reality. In *International Conference on Applied Human Factors and Ergonomics*. Springer, 119–126. doi:10.1007/978-3-030-79763-8_14

[75] Kevin Yu, Ulrich Eck, Frieder Pankratz, Marc Lazarovici, Dirk Wilhelm, and Nassir Navab. 2022. Duplicated reality for co-located augmented reality collaboration. *IEEE Transactions on Visualization and Computer Graphics* 28, 5 (2022), 2190–2200. doi:10.1109/TVCG.2022.3150520

Received 2025-07-25; accepted 2025-10-17