Untethered Displays: The Effects of Mixed Reality on Split-Attention in Fine-Motor Tasks

by

Griffin J. Hurt

Submitted to the Undergraduate Faculty of

the David C. Frederick Honors College in partial fulfillment

of the requirements for the degree of

Bachelor's of Philosophy

University of Pittsburgh

2024

UNIVERSITY OF PITTSBURGH

DAVID C. FREDERICK HONORS COLLEGE

This thesis was presented

by

Griffin J. Hurt

It was defended on

March 30, 2024

and approved by

Jacob Biehl, Department of Computer Science, University of Pittsburgh

Edward Andrews, Department of Neurological Surgery, University of Pittsburgh

Karim Kader, Department of Surgery, University of California, San Diego

Adam Lee, Department of Computer Science, University of Pittsburgh

Untethered Displays: The Effects of Mixed Reality on Split-Attention in Fine-Motor Tasks

Griffin J. Hurt, B.Phil

University of Pittsburgh, 2024

Augmented reality presents new affordances for users to interact with information without limitations of the physical world. For instance, a surgeon in the operating room may want to adjust a monitor without breaking the sterile field, or a military technician may need to reference information from an immovable screen. These environments currently use either large displays or personal tablet devices to present this information to the user. We have developed a system allowing users to segment a physical display, virtualize the segments as holograms in mixed reality, and position the elements around the task space. We also conducted a user study to understand the effects of display modality (mixed reality, tablet, or large display) on participant performance and preferences. We chose Lego construction as our experimental task, as it is reinforced by the literature and mirrors the critical interaction challenges that exist in constrained environments. We found no significant difference in performance between the three modalities, indicating mixed reality may provide a drop-in solution to places where displays or tablets are disadvantaged. The tablet modality was subject to familiarity bias, which may have affected participant rankings. Participants were likely to recommend mixed reality for hypothetical people completing tasks with more restrictions than the experimental task, regardless of their personal ranking order. Overall, these findings present design implications for replacing large displays and personal tablets with mixed reality in information-rich environments.

Table of Contents

List of Tables

List of Figures

Preface

I would like to thank Dr. Jacob Biehl and Dr. Edward Andrews for their continued support for my research and personal growth during my two years of undergraduate study at Pitt. I would also like to thank Dr. Karim Kader and Dr. Adam Lee for their time reviewing my thesis. Thanks to Talha Khan for his research and personal mentorship, Nico Kass for his assistance with the quantitative analysis, the Surreality Lab for their resources (both in knowledge and equipment), my parents and brother for shaping me into the person I am today, my roommates Joe Altvater and Carson Gollinger for keeping me focused and productive, and my partner Ashley Ingenito for providing me with support throughout my BPhil process.

1.0 Introduction

Since the beginning of the Information Age in the mid-20th century, purely computational jobs have widely disappeared in favor of electronic devices performing these tasks. As a consequence to this shift in the service industry, most careers today are knowledge-based and information-mediated, wherein workers must retrieve information from a source and use it to complete their tasks. This is especially true in collaborative environments. Although many methods exist for providing dynamic information in a visual manner using modern technology, most information-rich environments rely on either large displays or personal tablet devices to present data.

One example of an information-mediated collaborative task is manufacturing or processing plant control. These facilities require employees to view and understand large amounts of visual information and react quickly in the case of an emergency. This information is generally provided to employees by way of a personal tablet or mobile device. Serious consequences can occur when information presentation is not optimized in manufacturing plant control. In 1994, a petroleum cracking plant exploded when a flammable liquid was moved into a vessel that had a closed outlet. Human-computer interaction factors played a role in the accident, as the information displayed on control equipment was largely textual and colors were not displayed, so it was difficult to detect environmental changes [24].

Command and control centers are another example of an information-rich environment that requires first responders to consult real-time battlefield data and respond in a timely manner [2]. In modern command and control centers, a variety of tablet systems and large displays are used for information presentation [28]. Communication in command and control situations is made difficult by the presence of both unshared information, which only a select portion of the team has access to, and shared information, which is available to all. Previous work has indicated moving more unshared information into the common view may improve communication in command and control; however, increasing the amount of shared information raises cognitive load and may cause first responders to lose sight of important data [10].

Surgical operations are another information-mediated task. Surgeons must pay close attention to a variety of information sources, including monitors and alarms, while operating to ensure the patient remains stable [19]. In the operating room, given the need for sterility, interaction with information systems is often inefficient, which could have detrimental effects on surgeon comfort, surgical time, and ultimately patient outcomes. Surgeons are usually unable to interact with traditional input methods, such as keyboards and mice, in the operating room because they quickly become contaminated during surgery [6, 26]. Even if surgeons are presented with the opportunity to use these systems, the equipment is usually covered by a plastic sheath that must be changed, creating delays in a life-critical situation [6, 31]. Alternatively, a surgeon may ask another member of the surgical team to make changes to the machinery instead. However, interaction issues exist with communication in the operating room, as surgeons often give orders in a negative manner to surgical staff regarding equipment changes [29].

1.1 Spatial Computing in Constrained Environments

Evidently, there are a multitude of design limitations that exist within constrained environments like the operating room and command and control centers that prevent ergonomic interaction with traditional display modalities. Despite this, these environments still rely on these modalities to provide mission-critical data to workers, either due to the status quo or perceived lack of a better option. The last two decades have seen the adoption of spatial computing technologies increase exponentially with the development of technological advances. Spatial computing comprises three major display and interaction technologies: virtual reality (VR), augmented reality (AR), and mixed reality (MR). Novel visual and interactivity affordances are presented by spatial computing, including perceivable depth, personalization, environmentally anchored holographic objects, and hands-free interaction. This increased set of interaction modes may create a solution for the limitations that exist within information-rich environments. For instance, while a surgeon may not be able to physically move a large display for a better view while operating, they may be able to adjust a virtual monitor while keeping the sterile field or use eye gaze to select which monitors to prioritize.

1.1.1 Brief History of Spatial Computing

Virtual reality was the first spatial computing term to be coined, originating in 1989 by Jaron Lanier when attempting to describe the "Ultimate Display" that would mimic the real world in all sensible ways [1]. For nearly 25 years following its conception, the technology existed primarily in academic and industrial spheres with limited public interaction. However, in 2012, Oculus launched a campaign on Kickstarter to bring the technology to the public in the form of their headset the Oculus Rift. This began the "second wave" of virtual reality, which paved the way for future innovation in publicly accessible head-mounted display technology.

Concurrently with the creation and development of VR, augmented reality technology was being developed. The term was coined in 1990 by Tom Caudell and David Mizell, two Boeing engineers working on a head-mounted display to display plane schematics on the factory floor [5]. However, the concept of AR dates back to World War II, when the British Military developed a system to display radar information on the windshield of a fighter plane. Perhaps the most notable example of AR in recent public conscience is Pokémon Go, which was the first AR mobile game to reach the top of download charts [25]. The main distinction between VR technology and AR technology lies in how immersed the user is: virtual reality technology completely immerses a user in a virtual environment, while augmented reality simply provides the user with virtual elements.

While the term "mixed reality" is broad in its possible meanings, the term is generally used to refer to one or more of four things [30]:

- a mix of real and virtual objects in a single display modality on a spectrum between a fully real and fully virtual world, which includes VR itself
- a synonym for AR
- a combination of VR and AR together
- "strong" AR, with greater capabilities than a system like mobile AR

These descriptions are not mutually exclusive, and the true description of mixed reality likely lies within several of these categories. This work will refer to mixed reality under the first description, as the system designed for this study allowed users to see and interact with the physical world while utilizing virtual elements.

1.2 Summary of Contributions

The particular benefits that spatial computing offers to mediate information presentation in these environments lead to several interesting human-computer interaction research questions:

- **RQ1:** How does display modality impact users' ability to perform tasks that are aided by digital information?
- **RQ2:** How does display modality impact users' interaction behaviors with digital information?
- **RQ3:** What are users' perceptions and concerns about interacting with information across physical and virtual barriers?

To test these research questions, we designed and built a system that permits users to capture their screen contents, segment the contents of the screen into discrete portions, and transmit the visual information to a head-mounted display. The received video stream can then be presented as environmentally-anchored holographic windows within mixed reality. We also performed a study to measure the effects of three modalities representative of modern information presentation technology (large display, personal tablet, and mixed reality) on performance, task load, and situational awareness. The main contributions of this work lie in design implications for adding mixed reality to information-mediated tasks that currently use large displays or personal tablet technology.

2.0 Related Work

2.1 Mixed Reality in Task-Guided Work

Much previous work has explored the use of mixed reality for assistance in task-guided work. Daling and Schlittmeier analyzed 24 articles that detailed the use of mixed reality in manual assembly tasks [8]. They found that the mixed reality allowed information to be displayed in a way that was easier to understand through the overlaying of virtual objects and the customization capabilities of the technology allowed participants to work at their own pace. They also reviewed handheld information presentation options, like personal tablets, which significantly outperformed large screen-based solutions in satisfaction with performance, usability, and willingness to recommend the system. Tang et al. also explored the effectiveness of providing instructions in augmented reality for assembly task with performance and cognitive load measurements [32]. They found that task load, a measured defined by mental, physical, and temporal demands as well as effort, frustration, and performance [13], was significantly lower when participants were shown augmented reality instructions as opposed to instructions displayed on a large monitor. However, there was no significant difference in performance, defined as time of completion and accuracy of a Lego assembly task, between the two modalities. 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 [34]. 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. This demonstrates the benefits of a mixed-modality system in environments with both shared and personalized information.

2.2 Multi-Display Settings

Khan et al. explored the capacity for mixed reality to replace planar displays in the operating room at varying levels of latency by asking participants to lay 16 lines of suture in under 3 minutes [18]. They found no significant difference in performance between the monitor and mixed reality modalities when both had 100ms of latency, and no significant difference in task load (measured by the TLX) between the two. Pavanatto et al. tested how well holographic displays in mixed reality could replace physical monitors by having 18 participants grade undergraduate assignments using purely virtual, hybrid, and purely physical setups each with 3 screens [23]. 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. Medeiros et al. investigated the usefulness and configuration of holographic displays in shared transit spaces [21]. Given the space constraints of public transit, it is impossible to configure large physical displays, which motivates the usage of mixed reality to create virtual screens. They found that participants' decisions on how to place the holographic displays depended mostly on 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. also investigated the placement of holographic displays in airplane environments [22]. Participants stated that the ability to create multiple displays of varying sizes in mixed reality increased the ability for multitasking compared to singular displays that required tabbing. Participants also expressed differing preferences for configuring displays horizontally or vertically, shedding light on the benefits of mixed reality to provide personalized display setups.

Gauglitz et al. tested the effectiveness of remote collaboration using a tablet. Participants were asked to guide another participant through landing an airplane by showing specific control panel elements they needed to manipulate [11]. The content of the tablet was first with just streamed video from the other participant, then the video with static markers overlayed, and finally the video with world-anchored markers. They discovered that the world-anchored markers were preferred amongst participants compared to the streamed video and static markers. Ryall et al. investigated the effects of group size and table size when using a tabletop shared display [27]. They found that the size of the group impacted the manner in which the shared display was configured, and that information may need to be shown in multiple orientations or on multiple displays in large groups. Izadi et al. developed an interactive system to share media in a public environment called *Dynamo* [16]. A key feature of the system is the ability for users to move information from the shared context on the display to their personal device. When interviewed about *Dynamo*, participants found these transferring features to be beneficial for collaboration, demonstrating the importance of being able to prioritize shared artifacts by displaying them on a personalized device.

2.3 Information Rich Environments

An early study by Haines et al. investigated the behavioral effects of heads up display that showed relative plane information in the pilot's field of view [12]. They found that, by providing information through the heads up display, the pilot almost never diverted their eyes to the instrument panel, and decreased their total head and eye movement. Khan et al. explored the possibility of using mixed reality in the operating room and conducted a variety of feasibility studies [17]. Surgeons reported mixed reality holograms to be adequate in terms of color and resolution; however, the latency of the system (250 ms) was unacceptable for endoscopic views. Czerwinski et al. explored the effects of a wider field of view and gender bias in navigating virtual environments [7]. They found that increasing field of view increased participant preferences and appeared to reduce gender bias. They also noted that performance and preference are complex, and can be affected by a multitude of factors on the individual level.

2.4 Technology Acceptance Model

The Technology Acceptance Model (TAM) is a key model for understanding why and how information technology is integrated into users' personal and professional lives [20].

The theory was developed in the 1980s to help understand why workers might not effectively utilize information tools provided to them and has since been widely cited for understanding both general-purpose and healthcare-oriented information technology [14, 20].

The original model found that there were three key factors that affected technology acceptance: perceived usefulness, perceived ease of use, and attitude [9]. The three factors are not independent of one another, and all elements of a technology's design impact the component to some degree. Since its initial creation, the model has been expanded by several other models, including the Technology Acceptance Model 2 (TAM2) [35]. The second iteration considered additional factors including job relevance, social influence (called "subjective norm"), and result demonstrability.

Given that spatial computing technology is novel to a large portion of the population, TAM and TAM2 may shed light into how the technology will be accepted as it begins to be integrated in more environments.

3.0 Methods

3.1 Study Design

3.1.1 Task

To investigate our research questions, a suitable study task to induce split attention in an information-mediated environment was necessary. Besides just creating split attention, a preferable task would also present a large amount of information and require participants to perform spatial reasoning, visual search, sequencing, cognitive reduction, and bimanual manipulation, as these are critical interaction challenges within operating rooms, command and control centers, and factory control rooms. Prior work studying mixed reality's effects on assembly tasks has used Lego assembly as an analogue for machine maintenance and other fine-motor tasks [32, 4]. Given its validation by previous research, widespread understanding by the general public, and overlap with the previously detailed requirements, Lego set assembly was chosen as the task for this study. To increase the effects of split attention, we augmented the assembly task by requiring participants to switch between two Lego sets randomly at irregular breakpoints.

3.1.2 Participants

We enrolled 18 participants, the majority of which were undergraduate students studying computer science at University of Pittsburgh. Participants were recruited through in-class presentations and word of mouth. Participants were asked to complete a short demographics questionnaire before beginning the study, which included questions about gender, age, previous experience with AR on a scale from 1 to 5, with 1 being ”Little to None” and 5 being ”Extensive experience,” previous experience with VR on a scale from 1 to 5, with 1 being ”Little to None” and 5 being ”Extensive experience,” perceived value of AR/VR technology on a scale from 1 to 5, with 1 being ”Little to no value” and 5 being ”Significant value,” and vision capability (Appendix A.1). The inclusion criteria required participants to be over the age of 18, have English fluency, and normal vision or vision corrected to normal. After completing the demographics questionnaire, participants entered the study space and sat at a desk in front of a large screen, measuring 100 inches diagonally (Figure 2). Participants were then read the study instructions. Of the 18 participants, 11 were male, 7 were female, and 1 was non-binary. Participant ages ranged from 18 to 37, with the mean age being 22.28. Self-reported experience with AR ranged from 1 to 3, with a mean of 1.94. Self-reported experience with VR ranged from 1 to 4, with a mean of 2.50. Participants' perceived value of spatial computing technology ranged from 2 to 5, with a mean of 3.94 (see Figure 1).

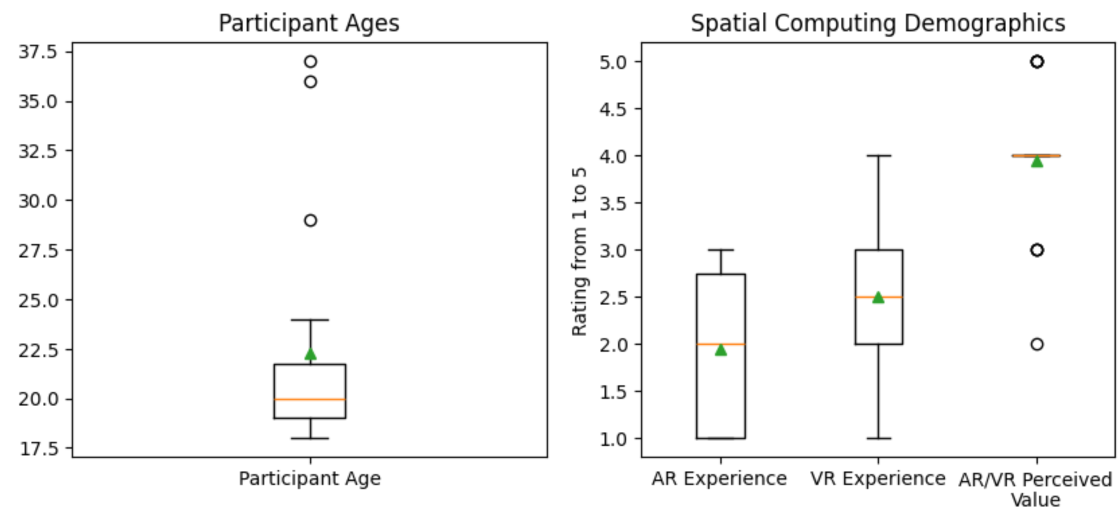

Figure 1: Participant demographic survey results.

3.1.3 Modalities

Each participant used three different display modalities to receive information about how to complete the construction task: large display, personal tablet, and mixed reality. The **large display modality** presented six segments on the screen in front of the participants (Figure 5). Five of the segments contained Lego set instructions, while the sixth segment contained the name of the active Lego set and the time remaining for the task. Two of the Lego instruction segments displayed the instructions for the sets the participant

Figure 2: The study space where participants assembled the Lego sets.

was constructing. The other 3 instruction segments were randomly chosen from a pool of other Lego instruction sets, including the instructions for other models in sets that could be constructed in multiple ways. These extraneous segments advanced at a randomized rate of one page per 1 to 10 seconds. The position of all segments was randomized.

For the **personal tablet modality**, participants were given an iPad device with an application that allowed them to select, position, and scale a subset of the segments from the large display on the tablet screen (Figure 3). Before beginning the task, participants were given time to familiarize themselves with the interface. Telemetry data was captured during the task measuring the number of frames in which the participant was interacting with the tablet.

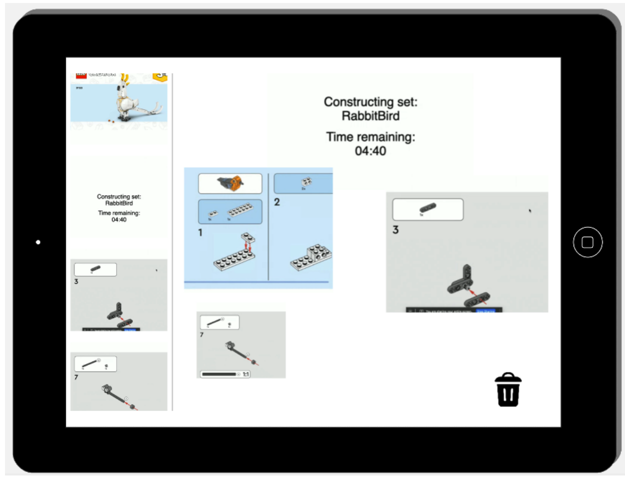

Figure 3: An example of the interface for the tablet modality.

Figure 4: An example of the interface for the mixed reality modality.

In the **mixed reality modality**, participants utilized a Microsoft HoloLens 2 to view the two instruction segments relevant to their task as holographic windows positioned above or near the corresponding Lego pieces (Figure 4). Before starting the task, participants were instructed on how to move, rotate, and scale the holographic windows using a pinching motion. Five of the participants positioned the holographic windows themselves. The remaining 13 participants had their holographic windows positioned by the research facilitator above the pieces corresponding to the relevant Lego pieces. Participants were permitted to adjust the segments if they wished before starting the task.

3.1.4 Procedure

For all modalities, participants advanced the active set instruction by verbally alerting the researcher. The order of the three modalities was counterbalanced to minimize the influence of extraneous factors, like carryover and ordering effects, on the experimental results. Each display modality was associated with two Lego sets, one deemed "difficult" and one deemed "easy" based on pilot performance data and piece count (Table 1). Of the six Lego sets, 4 were "3 in 1" sets in which the pieces could be assembled to make 3 different models. One of the "3 in 1" sets was a Lego Technic set, which has different piece types than a traditional Lego set. While using a modality, participants were forced to switch between construction of the two Lego sets at a random time interval between 15 seconds and 1 minute. The signal to switch was given as an auditory beep and reflected in the informational segment on the large display. Participants were afforded 5 minutes per modality.

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

Modality	Easy Set	Difficult Set
Large Display	Unicorn* (145 pcs.)	Shark* (230 pcs.)
Tablet	Taxi (124 pcs.)	Rabbit* (258 pcs.)
Mixed Reality	Nest (232 pcs.)	Dump Truck*† (177 pcs.)

*Indicates a "3-in-1" set †Indicates a Lego Technic set

Following timer expiration for each modality, the number of completed instruction pages for each set was recorded. Participants were asked to complete an experience questionnaire that measured task load, system usability, and situational awareness using the NASA-TLX [13], System Usability Score (SUS) [3], and Situation Awareness Rating Technique (SART) respectively [33]. After completing all three modalities, participants were asked a series of questions in a semi-structured interview (Appendix A.2). The first section of questions inquired about participant modality rankings, attention distribution, performance per modality, and situations in the participant's everyday life where the technology may be usable. The second section of questions asked participants to consider three hypothetical people and their possible applications of the modalities. The goal of these questions was to determine the task-specificity of participant responses. The first hypothetical person, named Bob, was detailed as an employee in a power plant who must reference a large monitor in his workspace. In his workspace, the large monitor is used collaboratively, such that every segment of the monitor is not useful to every technician. The second hypothetical person, named David, was described as a surgeon who must reference a large monitor with patient information and video feeds while maintaining a sterile environment that limits his ability to interact. The last hypothetical person, named Alice, was detailed as a car mechanic that must reference diagnostic data on a large monitor without dirtying the display by adjusting it. All three hypothetical people had jobs that placed them into a constrained enviorment, compared to the unconstrained environment of the experimental task. Participants were prompted one final time for additional thoughts on the modalities before conclusion of the study.

3.2 System Design

The MR Lego instruction delivery system comprises four parts: the signalling server, the streaming server application, the MR streaming client, and the instruction application.

The instruction application (Figure 5) was written in HTML, CSS, and JavaScript and designed to be run in a web browser due to the web's inherently visual interface. The instruction application was developed separately from the streaming software so that it could run independently. This functionality was implemented to create the Large Display modality, which does not require streaming. When the instruction application initializes in a web browser, the research team member selects the two Lego sets that the participant will be constructing, with the initial set selected as "Set 1". The user can then click the screen to make the application encompass the full screen and press the space bar to begin showing instructions and start the timer. The user can advance the currently active Lego set to the next instruction by pressing the space bar. After time expires, the user is brought to a new page that displays a counter with how many pages of instructions were completed on each set.

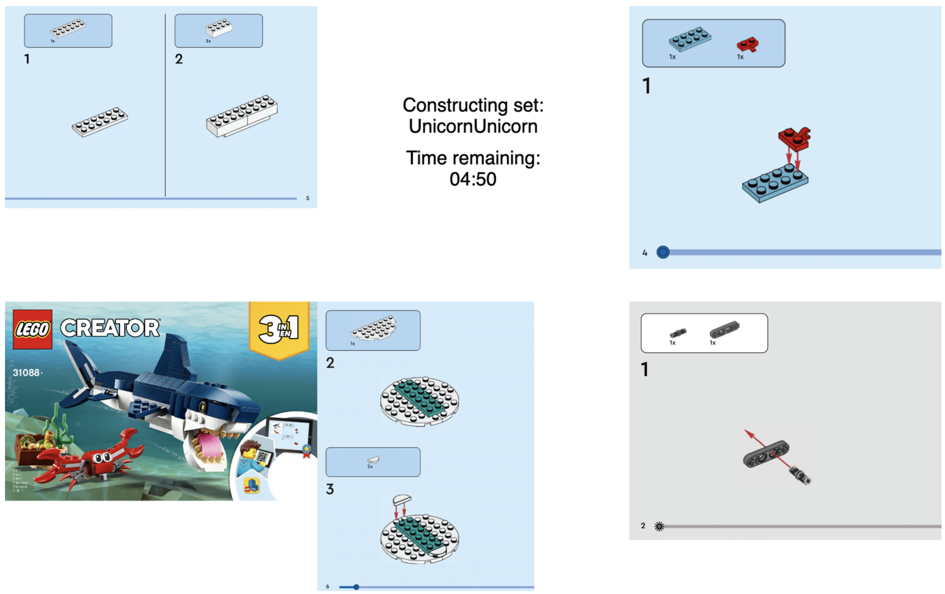

Figure 5: The instruction application with the shark and unicorn sets selected. The unicorn set is currently active.

The streaming server is also a web-based application developed in HTML, CSS, and JavaScript. The application can either stream a camera connected to the computer or a capture of the computer's screen, as seen in this study. The underlying streaming technology for the server is WebRTC, as it provided the lowest latency of the available streaming protocol options. WebRTC requires a signalling server that is shared between peers to exchange information about available communication channels (known as ICE candidates) and a connection string. In this system, the signalling server is implemented in NodeJS with the Socket.IO communication protocol. When starting a new streaming session, the user must first choose an input source for the stream (either webcam or screen capture) and provide a streaming resolution. For this study, the screen was captured at a resolution of 1280 pixels by 720 pixels. The user can then choose to segment the screen by drawing boxes around the displayed preview or select one of the default segmentation options (4 equally-sized segments or 6 equally-sized segments, but these could be expanded dependent on use case). Video is streamed to clients in a single stream; segmentation is performed on the client with bounds provided by the streaming server. After the display is configured and segmented, the user may start the server to begin streaming.

The MR system (Figure 4) used to deliver Lego instructions to participants was built primarily for the Microsoft HoloLens 2, a see-through MR headset. The Microsoft 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. This is in contrast to passthrough MR devices, which capture the outside world with cameras and reproject it within the headset. Additionally, the HoloLens 2 is considered to be a state-of-the-art device and is widely used in the literature [17, 18, 23]. The MR streaming client was built with the WebXR standard using the THREE.js JavaScript library. A web-based format for the application was chosen so the system could be used on multiple MR headsets without additional configuration and for better WebRTC support. While other commonly used MR development platforms like Unity do have support for WebRTC, WebXR required the least amount of configuration to connect to the streaming server. Once initialized, the MR system must first be calibrated by outlining the bounds of the large display. This is accomplished by the user pinching their fingers at the four corners of the display. Following calibration, the system connects to the signalling server and exchanges information with the streaming server to open a WebRTC connection. After establishing the WebRTC connection, the MR client displays a faint hologram over the large display as a guide, segmented based on the bounds provided by the streaming server. The user can choose which segments of the large display to virtualize by pointing their hand at a segment on the faint display guide hologram, pinching their fingers, and pulling away from the screen. This creates a holographic clone of the selected display segment. The user can move this segment by pinching and dragging with one hand or scaled/rotated by pinching and dragging with two hands. Only one clone of each display segment can be created at one time; if the user attempts to create a new version of an already cloned segment, the previously created clone is destroyed. Users may delete a virtual segment by placing it near its original position on the "ghost display". Virtual segments are redrawn from the WebRTC video stream at 60 frames per second, as dictated by the THREE.js library rendering pipeline.

The tablet system (Figure 3) is constructed similarly to the MR system. The same signalling server, streaming server application, and instruction application were used for both systems; however, it uses a separate streaming client designed specifically for the tablet. The tablet client is also packaged in a web application, allowing for cross-platform deployment. Upon initialization, the client connects to the signalling server to exchange WebRTC connection information with the streaming server. Once the WebRTC connection has been configured, the tablet client builds the left sidebar interface using the segmentation bounds provided by the streaming server. 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 larger area shown on the middle of the tablet screen. Once segment clones have been created in the "playspace," the user may manipulate them by dragging with one finger or scaling using pinching gestures with two fingers. 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 at once; 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 position the segment near the trash can object in the bottom right corner, which will cause it to be deleted. All segments, both on the sidebar and in the "playspace," are redrawn from the WebRTC video stream at 30 frames per second. If the user is actively interacting with a virtual segment during a frame, the interaction is logged and transmitted to the streaming server, where it can be downloaded by the user in JSON format.

Across all modalities, connection establishment follows this basic structure (Figure 6):

1. The user starts the signalling server.
2. The user starts the streaming server application, configures the video source with segmentation (if applicable), and starts the streaming server.
3. The user starts the client for a given modality.

4. The client connects to the signalling server, exchanges information about streaming capabilities, and provides a WebRTC connection candidate to the streaming server.
5. A WebRTC connection is established between the modality client and streaming server.
6. On each frame, the client performs segmentation of the incoming video stream using the segment information provided by the streaming server. Additional telemetry data may be provided dependent on modality.

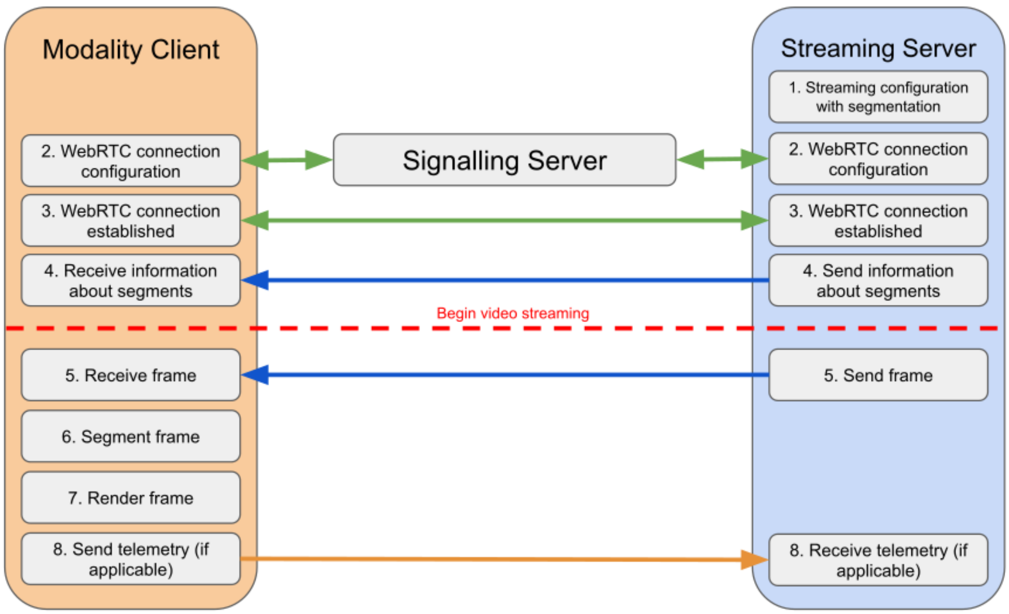

Figure 6: A diagram of streaming configuration and operation for the system.

4.0 Results

4.1 Metric Calculation

Participant performance was calculated per modality by summing the number of bricks placed for each of the two sets. Number of bricks was chosen as the performance metric, as it represents the smallest unit of task improvement for Lego set assembly. Total TLX 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. Total SUS was also calculated 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. The same method was used to sum total SART scores, with positive situational awareness 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.

4.2 Metric Summaries

The tablet modality saw the highest mean performance, with an average of 22.50 bricks between the two sets. The large display had the second highest mean performance, with an average of 22.44 bricks placed between the two sets. Mixed reality had the lowest performance on average, with 20.28 bricks placed. No significant difference in performance was observed between the three modalities (Figure 7).

The tablet modality had the lowest mean TLX score (lower scores represent lower task load), with an average score of 246.67. Mixed reality had the second lowest mean TLX score,

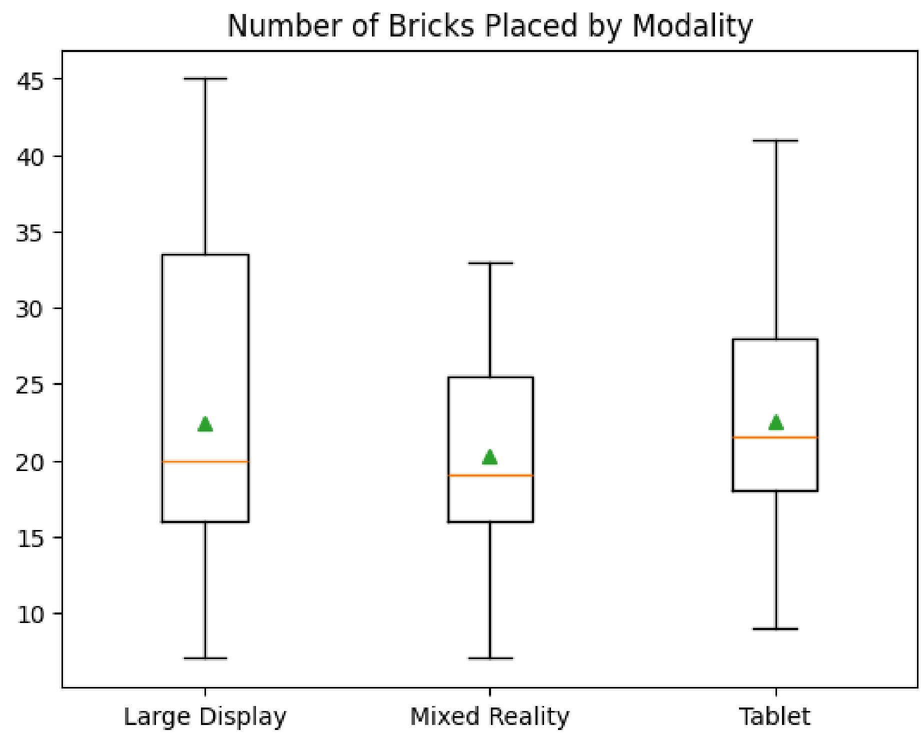

Figure 7: Performance, measured as number of Lego bricks placed, by modality as a box plot.

with an average of 260.00. The large display modality had the highest mean TLX score, at 264.44. For system usability, tablet had the highest mean SUS (higher values indicate better usability) at 43.67. Large display had a SUS mean of 39.06, and mixed reality had a mean SUS of 38.94. For situational awareness, lower SART scores represent greater situational awareness. The large display modality had an average SART of 22.94, the tablet modality had an average SART of 23.06, and mixed reality had an average SART of 24.5. There was no significant difference between the three modalities for TLX, SUS, or SART scores (Figure 8). The tablet modality had the top mean participant ranking at 1.89, with 8 participants ranking it first, 4 participants ranking it second, and 6 participants ranking it last. The display modality had the next best mean participant ranking at 2.00, with 6 participants ranking it first, 6 participants ranking it second, and 6 participants ranking it last. The mixed reality modality had the worst mean participant ranking at 2.11, with 4 participants ranking it first, 8 participants ranking it second, and 6 participants ranking it last.

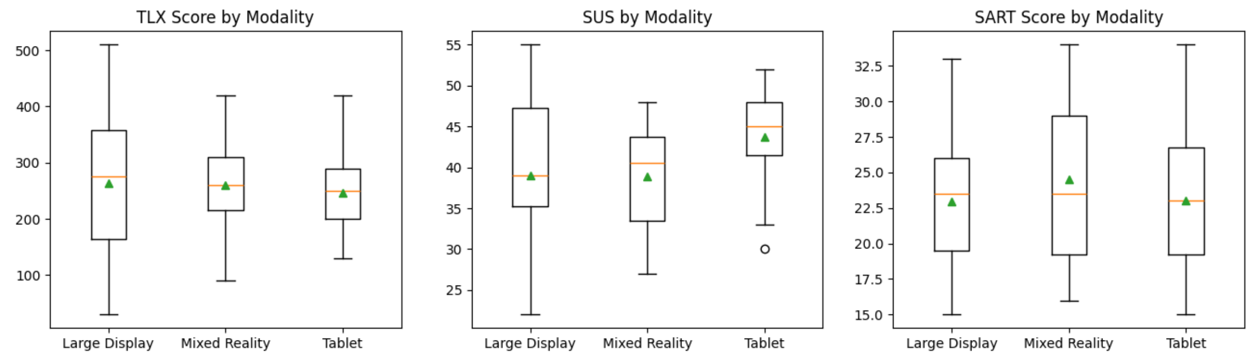

Figure 8: TLX, SUS, and SART scores by modality as a box plot.

Those who performed better on mixed reality than the other modalities ranked mixed reality significantly lower than those who did not ($p = 0.04$). Similarly, those who performed better on the large display modality than the others ranked the large display significantly lower than those who did not ($p = 0.04$). However, there was no significant difference in tablet ranking for those that performed better on the tablet than the other modalities (Figure 9).

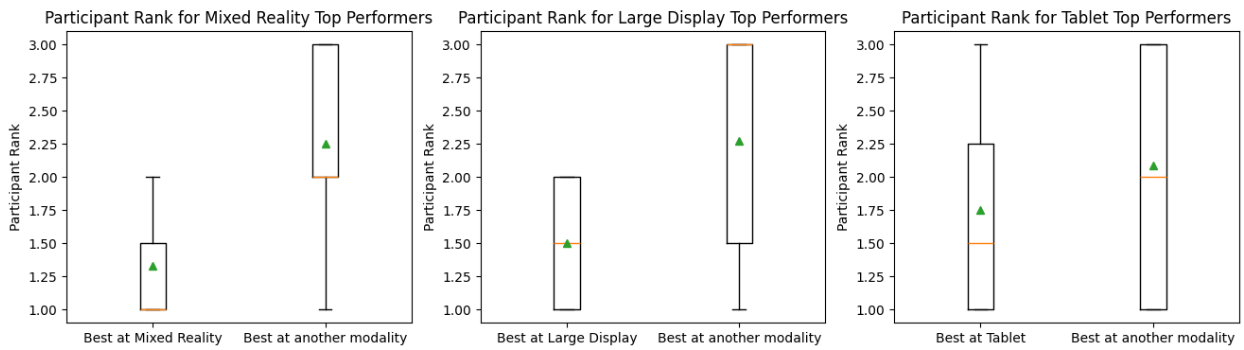

Figure 9: Participant ranking for the modalities they best performed in.

4.3 Correlations

The linear correlation between rank and performance and rank and SART were not found to be significant for all modalities. The linear correlation between rank and TLX and rank and SUS for the large display modality was found to be significant ($p = 0.01, m = 0.01, b = 0.80$ and $p = 0.01, m = 0.01, b = 1.07$ respectively), however not for the other modalities (Figure 10).

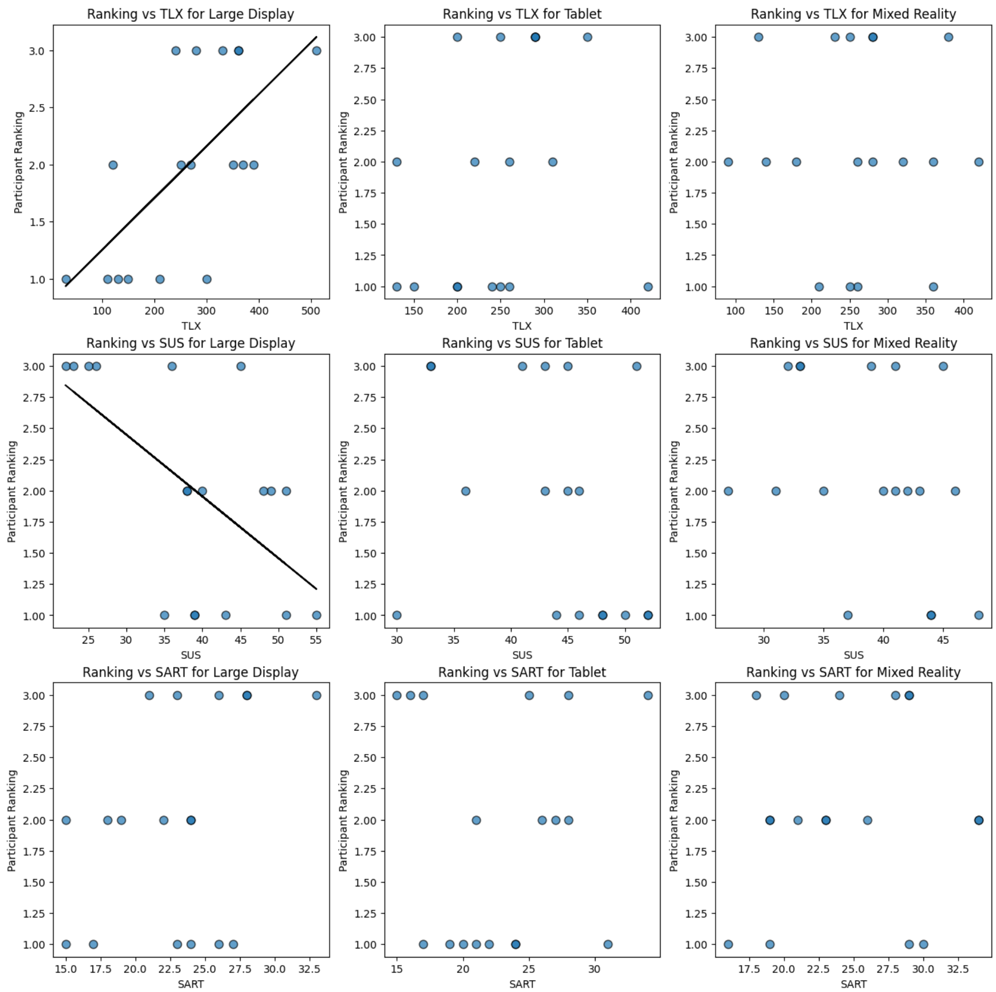

Figure 10: Correlation between ranking and TLX, SUS, and SART for each modality, with significant regression lines shown.

The linear correlation between situational awareness and performance (Figure 11) was found to be positive (lower SART values correspond to higher situational awareness) and significant for the large display ($p = 0.02, m = -1.34, b = 53.34$) and the mixed reality modalities ($p = 0.04, m = -0.59, b = 34.81$); however, the same relationship was not found to be significant for the tablet modality.

The linear correlation between tablet performance and number of tablet interaction frames (Figure 12) was found to be positive and significant ($p = 0.01$).

We found the linear correlation between performance for all three groups to be significant ($p = 0.01, m = 1.13, b = -0.39$ for large display versus mixed reality, $p = 0.01, m = 1.11, b = -2.55$ for tablet versus large display, $p = 0.01, m = 0.84, b = 5.50$ for mixed reality versus tablet) (Figure 13).

We did look for other correlations between rank, performance, previous mixed reality experience, and experience metrics (TLX, SUS, and SART); however, none were found to be significant.

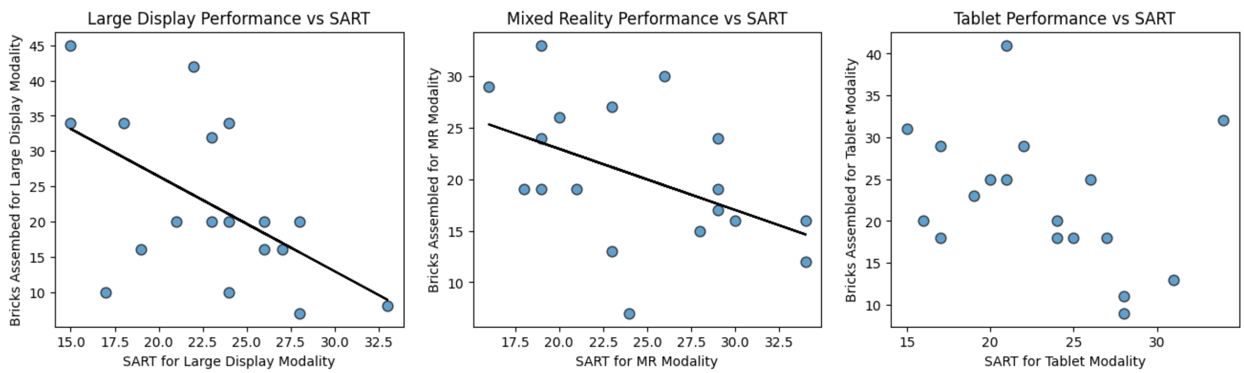

Figure 11: Correlation between SART and performance, with significant regression lines shown.

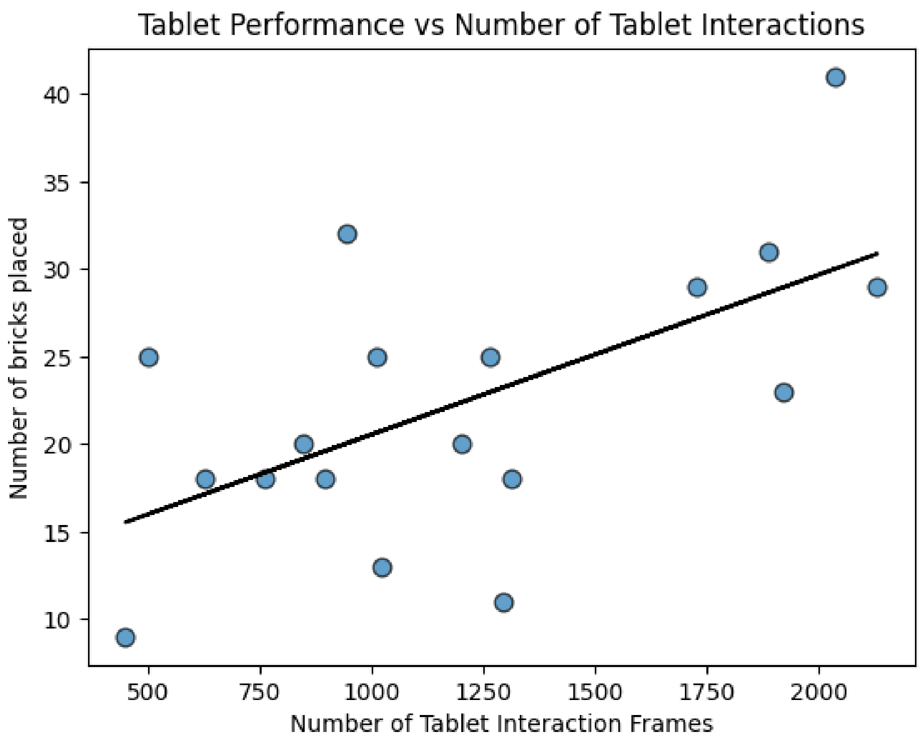

Figure 12: Significant correlation between number of tablet interaction frames and performance, with regression line shown.

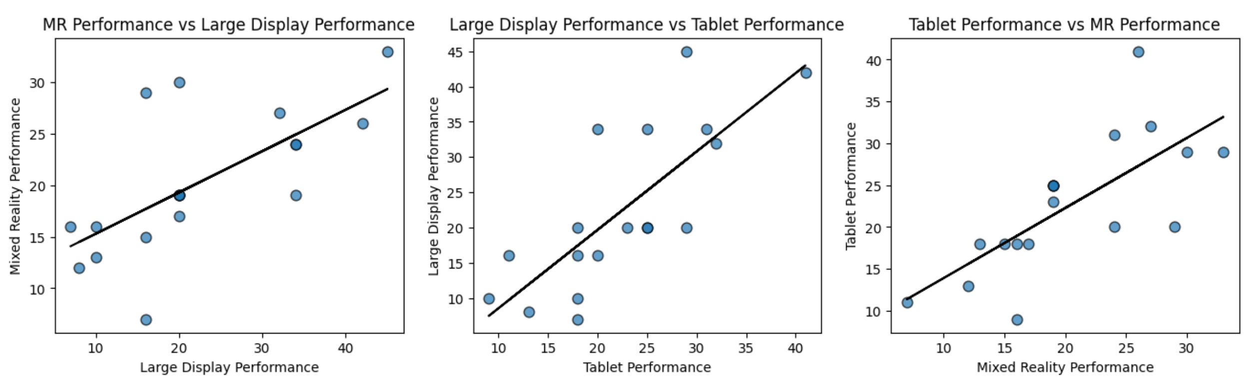

Figure 13: Performance correlations between all three modalities with regression lines shown.

4.4 Interview Responses

Table 2: Participant rankings by modality

Modality	Ranked First	Ranked Second	Ranked Third	Mean
Tablet	8	4	6	1.89
Large Display	6	6	6	2.00
Mixed Reality	4	8	6	2.11

Although all participants were asked for their rankings of the three modalities, given the format of semi-structured interviews, not every participant was asked every question. After ranking the three modalities, participants were asked whether they performed their rankings based on personal comfort or perceived task efficiency. The majority of participants (10) stated that their rankings were based on perceived task efficiency, 3 participants stated their rankings were based on personal comfort, and 2 participants stated their rankings were based on both equally.

All participants were asked which modality allowed them to best focus their attention. 6 participants stated that the large display allowed them to best focus their attention, 6 responded with the tablet modality, 5 participants responded with the mixed reality modality, and 1 participant responded that it was a tie between mixed reality and the large display. Participants that stated mixed reality helped them best focus their attention were asked whether the benefits of mixed reality outweighed the downsides of using the system, of which 4 responded with yes and 2 responded with no.

Researcher Do you think that the benefits afforded by the mixed reality system outweigh the burdens associated with it?

Participant That's a tough question because if it was just specifically for Lego sets, I think this is a little bit overkill to use a [head-mounted display] for this. But if this was something more serious, like maybe some sort of mechanical work or engineering... I definitely think that this could help... So for stuff like that, I do think the benefits would outweigh the burden...

12 participants were asked whether there were outside factors that impacted their performance besides the modality they used. 10 of these participants stated there were outside factors, while 2 stated there were not. Of the participants that stated there were outside factors, 7 stated the difference in difficulty between the Lego sets affected their performance, while the remaining 3 stated the main factor was becoming more comfortable with the task.

Researcher Do you think your performance on the task was only dictated by the configuration you were using or were there outside factors?

Participant Definitely not just purely the configuration. 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, especially in terms of colors and searching for pieces like the ones that were less homogenous in color...

The final portion of the semi-structured interview described three hypothetical people (Bob, a power plant technician, David, a surgeon, and Alice, a car mechanic) and asked which modalities would be most applicable to their jobs (Table 3). Each participant was asked about 2 of the hypothetical people, chosen at random. Of the 10 participants asked about the power plant technician, 4 stated he could use the mixed reality modality, 3 suggested he could use the tablet modality, and 3 suggested he could use the large display. Of the 14 participants asked about the surgeon, 13 stated he could use the mixed reality modality and 1 stated he could use the large display. Of the 12 participants asked about the car mechanic, 11 stated she could use the mixed reality modality and 1 stated she could use the large display. Interestingly, many participants that did not rank mixed reality first still stated that the hypothetical people could utilize the modality.

Table 3: Participant responses regarding the best modality for each hypothetical person.

Modality	Bob	David	Alice	Total
Mixed Reality	4	13	11	28
Large Display	3	1	1	5
Tablet	3	0	0	3
Total	10	14	12	36

Participant notes about the benefits and drawbacks for each of the three modalities were also recorded during interviews. The most commonly mentioned benefits of the large display included its comprehensiveness (5 participants), its minimized interaction (5 participants), its ease of use and perceived comfort (5 participants), its clarity (3 participants), and its capacity for collaboration (3 participants) (*"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 mixed reality 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."* - P3). The most common drawbacks of the large display included its physical size and external modification requirements (9 participants), its lack of personalization and poor information layout (5 participants), its increased cognitive load (4 participants), and its limited field of view and resultant movement requirements (3 participants) (*"Well, for example, the largest 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 AR instead."* - P9). Noted benefits of the tablet modality included its interactivity and personalization (11 participants), its proximity to the tasks (7 participants), its familiarity (6 participants), and its portability (4 participants) (*"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."* - P14). Specific drawbacks of the tablet included interaction issues or hand occupation (11 participants), its limited view capability (5 participants), and its physical space requirements (2 participants) (*"I had to keep like readjusting the screen and like zooming in and either like and then click on the next one and then I'd have to come back and forth between the two like tasks."* - P15). Mentioned benefits of the mixed reality device included its limited interaction and hands-free nature (14 participants), its personalization and increased task attention (12 participants), and its information detail, including depth capability (6 participants) (*"For example, the headset has the advantage that you have your two hands free and you don't have anything on the way."* - P16). Noted drawbacks of the mixed reality modality included system concerns (like battery life and cost) and physical discomfort (9 participants), image issues like color saturation (8 participants), lack of familiarity and interaction issues (8 participants), and obstruction of real-world view and lack of outside attention due to immersion (7 participants) (*"I feel like the mixed [reality] would have been my number one if you'd given me like 20 more minutes in it."* - P8).

5.0 Discussion

These findings lead to critical insights into design, preferences, and likely use styles or cases of multi-display collaborative hybrid environments.

5.1 Despite Lack of Preference, All Participants Believe Mixed Reality Has Value

Although many participants did not choose mixed reality as their preferred display type, almost every participant selected it as the best modality for surgeons and car mechanics. This sheds light on the task-dependent nature of information display replacement. In a task that is comparatively unconstrained, like power plant control, participant responses were split, mirroring participants' own preferences. However, in tasks that prevented workers from using their hands or otherwise limited their interaction with traditional display modalities, mixed reality emerged as the best modality. Evidently, participants were aware that the limitations of the environment play a substantial role in preferences for display replacement. This also reinforces the job relevance aspect of technology acceptance as detailed in the TAM2 [35].

5.2 Best Performance, Better Preference

For the mixed reality and large display types, participants that performed best with those modalities ranked them significantly higher. It is unclear whether this is simply a correlation or causative effect. Participants that performed better on a certain modality may have realized this while completing the task and subsequently preferred it due to its measurable benefit. Alternatively, if a participant preferred a modality outright, they may have been determined to perform better with it. The second interpretation suggest that, in real-world applications, providing users with their preferred modality would allow them to perform their best. In the future, a better understanding of this relationship must be developed to drive more performant conditions.

5.3 Situational Awareness and Performance

There was a significant positive correlation between situational awareness and performance for the large display and mixed reality modalities. This may point to the importance of mitigating split-attention for task-centered work. This finding appears obvious, as greater situational awareness would permit participants to focus more on their task, leading to better performance. However, it is not obvious why the tablet modality was found to not follow this trend. We discuss this unintuitive finding in more detail below.

5.4 Tablet Inconsistencies

The tablet modality is noticeably absent from several significant correlations between metrics. Although there are several possible explanations for this nonappearance, participant interview responses point to increased familiarity with the tablet modality. As explained by the TAM2, experience is a key factor in determining perceived usefulness, which in turn affects usage behavior [35]. Most participants stated in the interviews that they use a tablet in their every day life for a multitude of tasks including studying and note-taking; however, they do not commonly use a large display or mixed reality headset. This previous experience with tablet technology most likely impacted participants' preference for and performance using that modality.

Additionally, an increase in participant interactions with the tablet was correlated with a better performance on the modality. This is perhaps unintuitive, as it would be expected that increased interactions with the tablet would decrease task time, and consequently performance. However, similar to the conclusions from the relationship between best performance and preference, participant experience and preconceived notions may have impacted this metric. Those who commonly use or have unconscious bias toward the tablet may have utilized the modality's ability to bring information into and out of context in an optimal manner, allowing them to perform better. Altogether, this points to the fundamental differences between personalized information displays and shared information displays, as well as the familiarity bias that exists for the tablet modality.

5.5 The Road to Constrained Environments is Bumpy

This study provides substantial implications for utilizing mixed reality in environments previously occupied by large displays or personal tablets. There is a common assumption amongst proponents of mixed reality that the technology may be able to completely replace traditional displays in information-rich environments. Given that there was no significant difference in performance between the three modalities, mixed reality may offer a solution for information presentation in these environments. However, the preferences and performance of users completing a general task were not necessarily improved by replacing the display technology with mixed reality. This suggests that mixed reality has the most benefit to performance in specialized applications, but even so, not all concerns for expanding the use of mixed reality will be based around performance. Factors like familiarity, ergonomics, and acceptability will play a key role in determining how mixed reality is used.

In spaces like the operating room or command and control centers, there tends to be a fundamental tension between two factors: preferences and environmental constraints. Some people may prefer a certain information system for specific tasks; however, the limitations of their environment may prevent them from using their preferred modality. The technologies that lie between what is currently used and novel information displays will play a key role in bridging this divide. Luckily, places currently utilizing large displays or tablets for information presentation will not have their technology replaced overnight, resulting in a hybrid period where mixed reality is integrated with other solutions. Further work may benefit by investigating these hybrid setups combining mixed reality with other interactive display modalities.

6.0 Limitations and Future Work

Several limitations exist for this study. Only 18 participants were enrolled in the study due to time constraints and recruiting limitations. These participants were almost all college-aged students in the same geographic environment. There were preconceived notions about mixed reality among participants, including device cost and societal perceptions, which may have negatively impacted their rankings. Conversely, participants may have been subject to self-selection bias: those less familiar with mixed reality may have been less likely to enroll. However, most students were not experienced users with mixed reality.

As stated previously, the task that participants completed was contrived, abbreviated, and did not utilize many of the benefits of mixed reality. Participants directly mentioned legacy bias with the tablet modality, which indicates most participants were more familiar with using a tablet compared to a large display or mixed reality. In order to effectively study the differences between the modalities, participants would need to be experts in all three to prevent familiarity bias from affecting study results. However, it would be impossible to find a large group of users satisfying these conditions, so the task must be designed in a manner to cater to a diverse set of participant experience. This could be accomplished by modifying the abilities and interface of each modality dependent upon the results of a preliminary familiarity survey. For instance, participants that are experts with mixed reality would perform a more difficult task within that modality compared to the others. This study, however, explored only a single task without these modifications. It is also important to note that the task participants performed in this study did not utilize most of the benefits provided by mixed reality. The information was not provided with depth, the ergonomics of the large display was not a factor, and there was no requirement for collaboration. Our goal was to create an environment where no modalities were specifically disadvantaged, and every modality was in its best configuration. A comparison study that exercises the unique affordances provided by each modality is still needed.

There are plenty of limitations with the head-mounted display system as well. According to prior work, ideal AR 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 [15]. 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, which makes it far from perfect in providing a mixed reality experience.

6.1 Future Work

This study investigated participant preference and performance for different display modalities in a controlled lab environment. Ideally, these research questions would be studied in an authentic environment, like a simulated operating room or control center. We plan to conduct a similar study in the operating room using the same mixed reality system developed for this work. Surgeons will perform a normal endovascular surgery while the contents of the large display are recorded. Following the procedure, surgeons will enter a simulated operating room and utilize the system to segment and view the video recording of the surgical monitor in mixed reality while acting out the surgery. During the simulated surgery, surgeons will be asked to talk through their thought process, and will be asked semi-structured interview questions upon finishing the simulation. We hope that this study will shed light on the effects of an authentic environment on the research questions.

Future work may also focus on hybrid display approaches for interaction in constrained environments. We plan to conduct a future study in a simulated operating room investigating the benefits of mixed-modality information displays for surgical machine control. Surgical machines, like electrocautery devices, may need to be adjusted intraoperatively; however, the sterile environment of the operating room prevents direct surgeon interaction with the technology. Adjustment is usually accomplished by surgeons giving directions to a surgical technician verbally, who then makes the machine updates. However, as previously discussed, communication in the operating room can be difficult for a variety of reasons [29], which can hinder this adjustment process. In an attempt to remedy this, we plan to implement a system that allows surgeons to provide instructions using a mixed reality interface that mirrors the controls of the machine. The surgical technician will then use a tablet to view the adjustment instructions and make them accordingly. We hope this study will provide insight into the benefits of multi-modal and cross-modality information presentation in surgical environments.

7.0 Conclusion

The work of this thesis investigates the effects of different display modalities (mixed reality, tablet, and large display) used for information presentation on preference and performance within an information-mediated task. Through the development of a mixed reality system allowing users to segment physical displays and virtualize segments as holograms, we aimed to address the challenges faced by professionals in information-heavy environments. Although we found no significant difference in performance between the three display modalities, the qualitative results from participant interviews shed light into the complexity that underlies replacing traditional, physical methods for information presentation with mixed reality technology. These findings may inform design decisions for maximizing efficiency and comfort within constrained environments and help guide the integration of mixed reality for information presentation.

Appendix

A.1 Demographics Questionnaire

1. What is your gender?
2. What is your age?
3. What is your highest degree earned?
   - Grade School
   - High School Diploma
   - Bachelor's Degree
   - Master's Degree
   - Doctoral Degree
4. Do you have previous experience with the following?
   - Augmented Reality (1 being "Little to None", 5 being "Extensive experience")
   - Virtual Reality (1 being "Little to None", 5 being "Extensive experience")
5. What is your perceived value of augmented/virtual reality technology? (1 being "Little to no value", 5 being "Significant value")
6. Do you need to wear contact lenses?
7. Do you need to wear glasses?

A.2 Semi-Structured Interview Questions

1. Can you rank the configurations of the device (large display, personal tablet, mixed reality system) from best to worst?
   - **Probes:** Why did you place *[configuration]* at that point in the ranking? Was task efficiency or personal comfort more important to you in these rankings? What

aspects of *[configuration]* did you like or find useful? What aspects of *[configuration]* did you not like or hindered you? Which systems were you most and least comfortable using?

2. Which configuration allowed you to best focus your attention on the current task and why?
   - **Probes:** What aspects of that configuration made it better for focusing your attention? Why were those aspects important? Do you think that being able to focus more on the current task improved your performance? How did focusing your attention impact your cognitive load during the task?
3. Do you think that the benefits afforded by the mixed reality system outweighed the burdens associated with it, like wearing the headset and possible eyestrain?
   - *This question is only asked if the mixed reality system was chosen as the answer to the previous question.*
   - **Probes:** On your form you indicated that *[sentiment about AR/VR technology]*. Do you think your sentiment about virtual or mixed reality affects the way you view this configuration? *(If participant has previous experience with AR/VR)* Do you think your previous experienced with virtual or mixed reality technology affects the way you view this configuration?
4. According to the data, the configuration that you performed best with was *[configuration with best performance]*. Do you have any idea why that might be?
   - **Probes:** Do you think your performance on the task was only dictated by the configuration you were using or were there outside factors? *(If not chosen as best configuration)* You did not choose that as the best configuration. Why do you think you performed better with that configuration even though it was not ideal?
5. Can you think of any situations where one of these system configurations would be beneficial for you in your everyday life?
   - **Probes:** What specific elements of that configuration would make it useful for that task? Do you think the other configurations would also work for this task? Would any elements of the configuration hinder you during the task?

6. Can you consider any scenarios when one of these configurations would be better than another?
   - **Probes:** In that scenario, what elements of the configuration make it better than others? Would there be any foreseeable problems with that configuration for that specific scenario? Based on your description of that scenario, would you change your rankings at all?
7. Consider a hypothetical person Bob who works in a power plant. Bob needs to reference a large monitor in his workspace that gives him information about power consumption and what systems are available for him to send power to different locations. However, Bob works with many other technicians in his workspace who all reference the same monitor. Not every section of the monitor is useful to every technician. Do you think Bob could utilize one of the presented configurations?
   - **Probes:** Which configuration do you think would be most beneficial to Bob? What elements of this configuration are useful to Bob? Which elements would be a hinderance? Considering that scenario, would you change your ranking of the three configurations?
8. Imagine a hypothetical surgeon, David, who is operating on a patient. During the surgery, Bob must reference a large monitor with patient information and video feeds of scopes. Several other members of the surgical team also need to reference this information to provide care for the patient. David is unable to break the sterile field by touching the monitor to adjust it, and he is actively performing surgery with his hands. Do you think David could utilize one of the presented configurations?
   - **Probes:** Which configuration do you think would be most beneficial to David? What elements of this configuration are useful to David? Which elements would be a hinderance? Considering that scenario, would you change your ranking of the three configurations?
9. Consider a hypothetical person, Alice, who works as a car mechanic. Part of her job is referencing a large monitor with car data and diagnostic information. Her hands are always busy changing the car and she doesn't want to get grease on the monitor by adjusting it. Do you think Alice could utilize one of the presented configurations?

- **Probes:** Which configuration do you think would be most beneficial to Alice? What elements of this configuration are useful to Alice? Which elements would be a hinderance? Considering that scenario, would you change your ranking of the three configurations?

Bibliography

[1] Christoph Anthes, Rubén Jesús García-Hernández, Markus Wiedemann, and Dieter Kranzlmüller. State of the art of virtual reality technology. In *2016 IEEE Aerospace Conference*, pages 1–19. IEEE, 2016.

[2] Menelaos Bakopoulos, Sofia Tsekeridou, Eri Giannaka, Zheng-Hua Tan, and Ramjee Prasad. Command & control: Information merging, selective visualization and decision support for emergency handling. In *ISCRAM*, 2011.

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

[4] João Barroso, Luís Fonseca, Bernardo Marques, Paulo Dias, and B Sousa Santos. Remote collaboration using mixed reality: Exploring a shared model approach through different interaction methods. In *European Virtual Conference on Computer-Supported Cooperative Work*, pages 1–4, 2020.

[5] Donna R Berryman. Augmented reality: a review. *Medical Reference Services Quarterly*, 31(2):212–218, 2012.

[6] David Black, Michael Unger, Nele Fischer, Ron Kikinis, Horst Hahn, Thomas Neumuth, and Bernhard Glaser. Auditory display as feedback for a novel eye-tracking system for sterile operating room interaction. *International Journal of Computer Assisted Radiology and Surgery*, 13:37–45, 2018.

[7] Mary Czerwinski, Desney S Tan, and George G Robertson. Women take a wider view. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 195–202, 2002.

[8] Lea M Daling and Sabine J Schlittmeier. 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):589–626, 2024.

[9] Fred D Davis. *A technology acceptance model for empirically testing new end-user information systems: Theory and results.* PhD thesis, Massachusetts Institute of Technology, 1985.

[10] R. A. Fleming and J. L. Kaiwi. The problem of unshared information in group decision-making: A summary of research and the implications for command and control. Technical report, 2002.

[11] Steffen Gauglitz, Cha Lee, Matthew Turk, and Tobias Höllerer. Integrating the physical environment into mobile remote collaboration. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 241–250, 2012.

[12] Richard F Haines, Edith Fischer, and Toni A Price. Head-up transition behavior of pilots with and without head-up display in simulated low-visibility approaches. Technical report, 1980.

[13] Sandra G Hart and Lowell E Staveland. Development of NASA-TLX (task load index): Results of empirical and theoretical research. In *Advances in Psychology*, volume 52, pages 139–183. Elsevier, 1988.

[14] Richard J Holden and Ben-Tzion Karsh. The technology acceptance model: its past and its future in health care. *Journal of Biomedical Informatics*, 43(1):159–172, 2010.

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

[16] Shahram Izadi, Harry Brignull, Tom Rodden, Yvonne Rogers, and Mia Underwood. 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*, pages 159–168, 2003.

[17] 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. 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*, pages 267–270, 2022.

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

[19] Nicholas Kiefer and Andreas Hoeft. Display of information in the operating room. *Current Opinion in Anesthesiology*, 23(6):772–777, 2010.

[20] Nikola Marangunić and Andrina Granić. Technology acceptance model: a literature review from 1986 to 2013. *Universal Access in the Information Society*, 14:81–95, 2015.

[21] Daniel Medeiros, Mark McGill, Alexander Ng, Robert McDermid, Nadia Pantidi, Julie Williamson, and Stephen Brewster. 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):3640–3650, 2022.

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

[23] Leonardo Pavanatto, Chris North, Doug A Bowman, Carmen Badea, and Richard Stoakley. 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)*, pages 759–767. IEEE, 2021.

[24] Raghavendra Rao Polakonda. Information presentation on mobile device for plant operations. 2008.

[25] Philipp A Rauschnabel, Alexander Rossmann, and M Claudia tom Dieck. An adoption framework for mobile augmented reality games: The case of Pokémon Go. *Computers in Human Behavior*, 76:276–286, 2017.

[26] William A Rutala, Matthew S White, Maria F Gergen, and David J Weber. Bacterial contamination of keyboards: efficacy and functional impact of disinfectants. *Infection Control & Hospital Epidemiology*, 27(4):372–377, 2006.

[27] Kathy Ryall, Clifton Forlines, Chia Shen, and Meredith Ringel Morris. Exploring the effects of group size and table size on interactions with tabletop shared-display groupware. In *Proceedings of the 2004 ACM Conference on Computer Supported Cooperative Work*, pages 284–293, 2004.

[28] Laurel C Sadler, Somiya Metu, and US Army Research Laboratory Adelphi United States. *Intelligent Command and Control Demonstration Setup and Presentation Instructions*. US Army Research Laboratory, 2017.

[29] Sisilie Havnås Skråmm, Inger Lise Smith Jacobsen, and Ingrid Hanssen. Communication as a non-technical skill in the operating room: A qualitative study. *Nursing Open*, 8(4):1822–1828, 2021.

[30] Maximilian Speicher, Brian D Hall, and Michael Nebeling. What is mixed reality? In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–15, 2019.

[31] Emily Walker Stockert and Alexander Langerman. Assessing the magnitude and costs of intraoperative inefficiencies attributable to surgical instrument trays. *Journal of the American College of Surgeons*, 219(4):646–655, 2014.

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

[33] Richard M Taylor. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness*, pages 111–128. Routledge, 2017.

[34] Balasaravanan Thoravi Kumaravel, Cuong Nguyen, Stephen DiVerdi, and Bjoern Hartmann. TransceiVR: Bridging asymmetrical communication between VR users and external collaborators. In *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, pages 182–195, 2020.

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