NAIST-IS-MT1651204

Master Thesis

Is Spatial Memory Better Trained in Augmented or Virtual Reality?

Nicko Reginio Caluya

September 14, 2018

Graduate School of Information Science
Nara Institute of Science and Technology

A Master Thesis
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Master of ENGINEERING

Nicko Reginio Caluya

Thesis Committee:

Professor Hirokazu Kato	(Supervisor)
Professor Kiyoshi Kiyokawa	(Co-supervisor)
Associate Professor Christian Sandor	(Co-supervisor)
Assistant Professor Alexander Plopski	(Co-supervisor)

Is Spatial Memory Better Trained in Augmented or Virtual Reality?*

Nicko Reginio Caluya

Abstract

Work space simulations help trainees acquire skills necessary to perform their tasks efficiently without disrupting the workflow, forgetting important steps during a procedure, or the location of important information. This training can be conducted in Augmented and Virtual Reality (AR, VR) to enhance its effectiveness and speed. When the skills are transferred to the actual application, it is referred to as positive training transfer.

However, thus far, it is unclear which training enviroment, AR or VR, achieves better results in terms of positive training transfer. To compare the effectiveness of AR and VR for spatial memory training in a control-room scenario, I conducted a user study with 16 participants who had to memorize the location of buttons and information displays in their surroundings. A follow-up evaluation investigated the impact the training had on short-term and long-term memory.

Results of this study show that VR outperformed AR when tested in the same medium after the training. In a memory transfer test conducted two days later AR outperformed VR. My findings have implications on the design of future training scenarios and applications.

Keywords:

augmented reality, virtual reality, head-mounted display, training, spatial memory

*Master Thesis, Graduate School of Information Science,
Nara Institute of Science and Technology, NAIST-IS-MT1651204, September 14, 2018.

Contents

List of Figures

List of Tables

1 Introduction

Worker training is a costly and time-consuming process. It involves learning a variety of steps that must be performed for each task, as well as memorizing important locations. For example workers in a control room must memorize all elements of the various control panels. In industry and manufacturing, the working environment often contains a large variety of objects placed in different locations within a large space. Being able to quickly locate these objects is crucial for efficient task performance. As a result, researchers have developed various guidance systems in order to help users quickly locate specific objects and to build a spatial memory map of the environment [15, 17, 27].

1.1 Training Transfer

Making simulations of workers' tasks and conditions are helpful in the training process to enhance skills for actual work scenarios. In the case of operating a plane, pilots need to memorize locations of buttons and levers in order to maneuver the plane efficiently.

Many studies in industrial and organizational (I/O) psychology, a subset of applied psychology, define positive training transfer as "the degree to which trainees effectively apply the knowledge, skills, and attitudes gained in a training context to the job" [2]. In this case of cockpit training, the skill learned should be the simultaneous navigation of the interface and outside conditions. Baldwin and Ford visualized this training transfer into a model where work environment support and training content are important training inputs (Fig. 1.1).

After conducting an integrated literature review, Burke and Hutchins [4] concluded that further work still needs to be done in different aspects of the training, particularly in training environments. They suggested to use more systemic

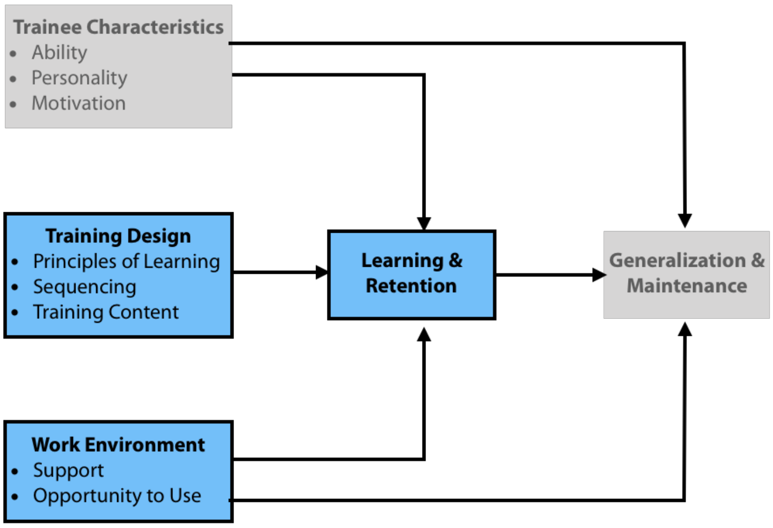


Figure 1.1: A model of the transfer process adapted from Baldwin and Ford [2], with highlighted components representing topics explored in this study.

approaches, rather than linear ones, when measuring training transfer. This study compares two different modes of training in simulated digital environments: Augmented and Virtual Reality (AR, VR). Traditionally, trainees experience workspace simulations exclusively in the real world. The perception of the trainee changes when employing VR and AR applications as training tools. The trainee views visual elements aside from viewing the real world, like the presence of virtual objects (or in the most extreme case, completely replacing the real world with a virtual one). Aside from these conditions that affect the user's perception, this study also looks at visual and proximal factors within training simulations. This study traces how the combined factors of training environment condition, proximity of locations, and visual Fsaliency of objects affect the process of memorizing spatial layouts.

1.2 Comparing Augmented and Virtual Reality

Investigating the differences between AR and VR starts with the definitions of the two. Milgram and Kishino [23] provided a taxonomy where AR and VR technologies fall in different parts of a virtuality continuum that ranges from a completely virtual environment to a completely real environment. Formal definitions for AR studies quote Azuma's work [1]. In his definition, AR "supplements the real world with virtual (computer generated) objects that appear to coexist in the same space as the real world". He also mentioned that three essential things constitute AR: such applications combine real and virtual (1), are interactive in real time (2), and are registered in three-dimensional space (3).

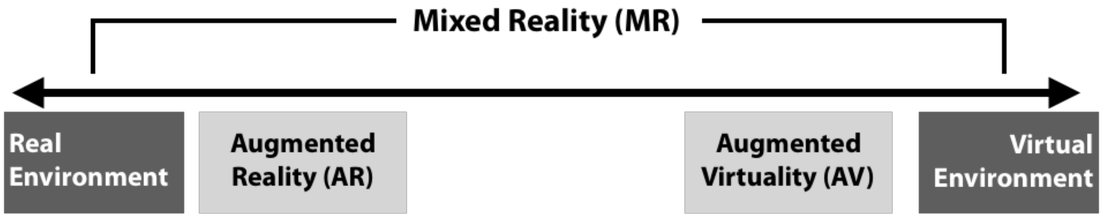


Figure 1.2: The virtuality continuum, adapted from Milgram and Kishino [23].

In this study, the VR condition is the absence of the real world or real objects in the user's view, effectively showing only virtual objects and environment. It is important, then, to know whether the real world is viewed partially or totally, to qualify if a certain environment is in AR or VR. This distinction also limits the study into observing only visual perception, and ensures there are no other cues that address other senses (e.g. audio and tactile feedback).

To further clarify the AR condition for this experiment, I would like to point out two cases of environments in the real world where an AR application can be deployed. On one hand, there are training environments which are not similar in appearance (e.g. size, realism) with the actual environment, even if they can be perceived in the real world. For example, miniature versions of cities or locales could be used for flood simulations to train urban planners or disaster response management teams, without needing the full scale and experience of an actual flood in an actual location. On the other hand, there are training environments that match the actual environment. An example for this is a flight simulator. Pilot trainees use flight simulators with almost the same experience as flying an actual airplane. The interactions and sensory experiences with the cockpit interface in the training environment is almost the same as the one in an actual airplane. In this study, I used the latter case, where the view of the training environment in the AR condition is the same as the actual environment used in the transfer.

In the domain of industrial applications, AR and VR can assist the training process by presenting guidance to trainees, or simulating a variety of situations that could not be presented in reality. [3, 5, 24]. Furthermore, AR and VR takes advantage of displaying virtual objects and providing users with more interactivity than by passively reading printed manuals or monitors. However, this study shifts away from a comparison between classic and digital techniques, the former being paper-based instructions and the latter being applications running on AR or VR devices. Instead, the comparison looks exclusively at applications in the virtuality continuum (Fig. 1.2), particularly at the differences between the training conditions under AR or VR. There have been studies that involve comparing skill performance in AR and VR conditions. In many cases, AR and VR could be applied for the same training scenario. For example, Dünser et al. [10] compared the effect of AR and a desktop computer-aided design (CAD) application on improving students' spatial understanding. They investigated this by using four tasks (folding, cutting, rotation, orientation). Although they found that training in AR did not outperform the CAD application, it does not mean that AR and VR cannot assist in spatial memory training, which involves a larger area. While the conditions presented in the study highlight the differences between AR and VR, they did not use the same device. In the CAD application condition, participants used traditional computer screens and interaction devices. In this experiment, I used only one device, an optical see-through head-mounted displays (OST-HMD), the Microsoft HoloLens, in the comparison between the VR and AR conditions.

While the choice of the display device can be a completely different factor to investigate, it remains an open question which training environment would be better suited for a skill like memorizing spatial maps. For these workspace simulations developed as an application in AR or VR, it is unclear to what degree the learned content can be applied in practical situations. Furthermore, it also unclear how different VR training is from training in AR. The experiment section highlights how the definitions of AR and VR inform the design choices of the system.

In AR, trainees experience the training scenarios in the target environment. One would expect AR training to have efficient skill transfer with a high retention rate from training to practical application. The main drawback of AR training systems is that the user must be present in the target environment. On the other hand, VR training allows users to train independent of their location, even from their homes. However, I expect VR training to have a lower retention rate of the information as the training conditions differ from the actual environment.

1.3 Research Approach

To better understand how well skills obtained in VR and AR are transferred to the actual application, this study focuses on an experiment involving a spatial memorization task in a control-room scenario (Fig. 1.3). Upon instruction, participants memorized the location of elements associated with labels in AR or VR, and then answered a short-term memory test immediately after each training session. After four rounds of training and short-term memory testing, participants came back 2 days after. Aside from short-term memory tests, a memory transfer test conducted 2 days after the training evaluated how well the learned material can be applied in a practical scenario. The results show that although VR outperforms AR training in the immediate post-training test, it performs worse in the memory transfer test. The results of the experiment suggest that AR should be preferred over VR for spatial memory training. At the same time, I expect VR to perform similarly well if the training is repeated over an extended period of time.

The main contribution of this study is being the first to perform a comparison between the effectiveness of AR and VR in transferring a spatial map from a simulation to the actual work environment. While the primary aim clarifies differences between AR and VR training through a comparison, this study also aims to provide design considerations when designing and deploying training applications for each condition (AR, VR).

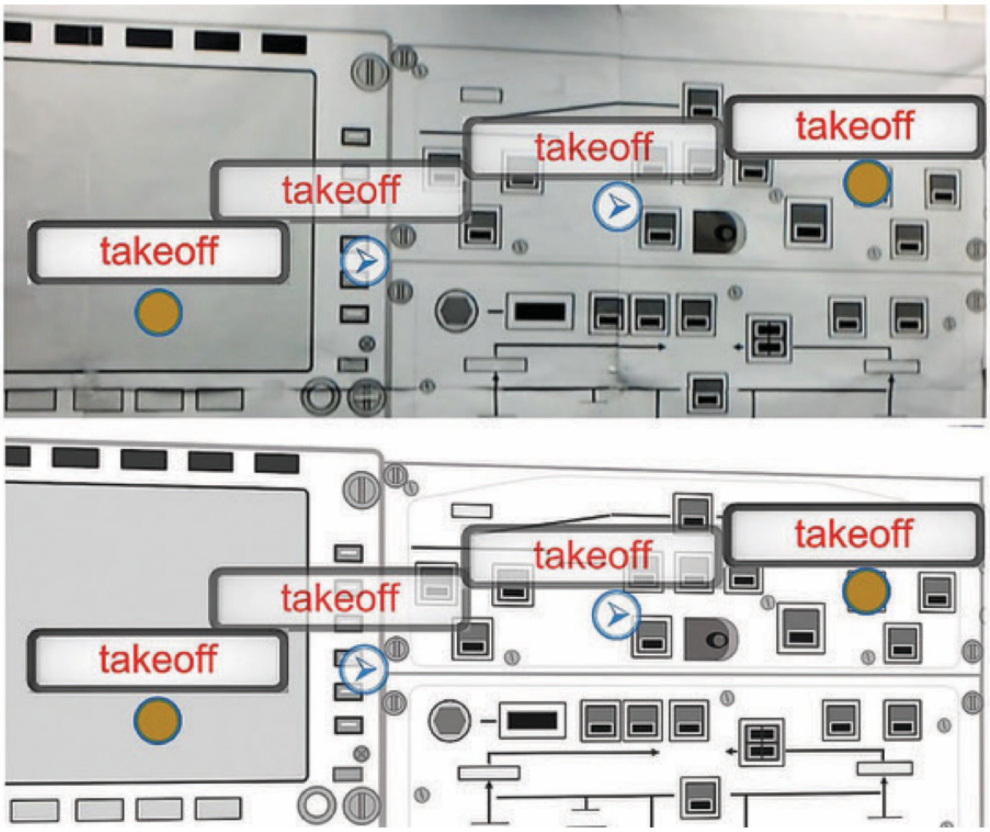


Figure 1.3: An illustration of the experiment task mechanics. The user wears a Microsoft HoloLens while performing a guided memorization task: an icon moves along with a word label ("takeoff") from an initialization area to a destination target. This task is done in two conditions: in AR (top) and in VR (bottom).

2 Related Work

In this section, I discuss related work that influenced the design and development of this experiment. Initially, this section defines what spatial memory is, then, looks at experiments and techniques that discuss how the environment design can assist spatial memory. Finally, an overview of methods and systems that were developed for spatial memory training shows how similar or different they might be from the system used in this study.

2.1 Spatial Memory

Spatial memory refers to the part of human cognition that makes us capable of retaining information about the geographic layout of an environment. Users with a spatial map of the environment can perform tasks more efficiently [6].

Scarr et al. [30] distinguish between two types of tasks associated with spatial memory: navigating through environments and remembering object locations. Users can learn spatial information either automatically or through effortful learning [16], i.e., some spatial information requires conscious effort from the user to retain it.

2.2 Approaches for Spatial Memory Support

There are two possible approaches to support spatial memory: present spatial information in a way that supports automatic learning, or create methods and systems that encourage effortful learning.

On one hand, various design factors can facilitate automatic learning of spatial maps. When we learn locations of objects in an environment, we tend to encode them relative to a spatial frame of reference [30], e.g., grids and landmarks.

Leifert [22] investigated if grid-based structuring and grid lines have an effect on spatial and content memory (Fig. 2.1). Her results show that while grid structuring has a positive influence on remembering the location of items, regardless of the presence of grid lines, it does not influence content memory. Furthermore, the presence of grid-lines yields worse performance on content memory.

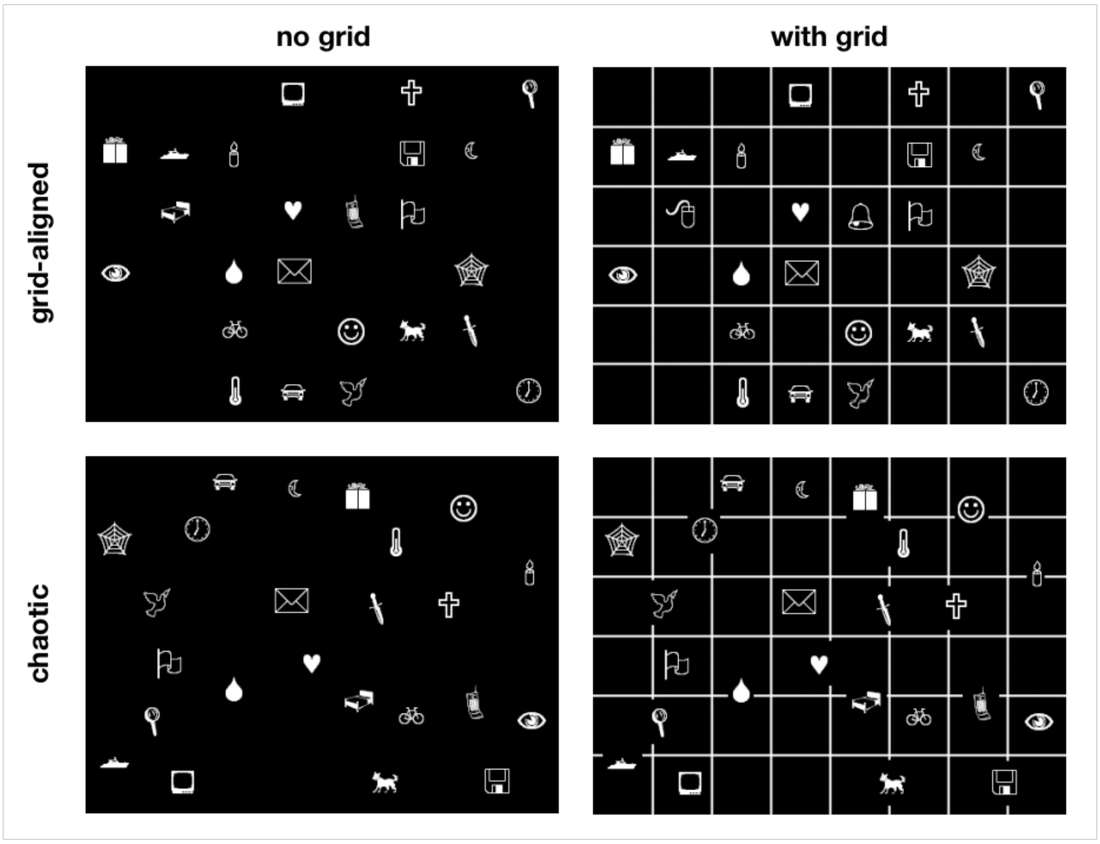


Figure 2.1: Matrix of canvas views showing arrangement of icon objects (grid-aligned, chaotic) and presence of grids (no grid, with grid), taken from Leifert's [22] experiment.

Uddin et al. [32] investigated if embedding artificial landmarks in grid-based user interfaces can positively influence spatial memory 2.2. They tested their hypotheses on grids with a small (64 items), medium (96 items), and large (160 items) number of items. Their results show that artificial landmarks improve spatial memorization in medium to large grids.

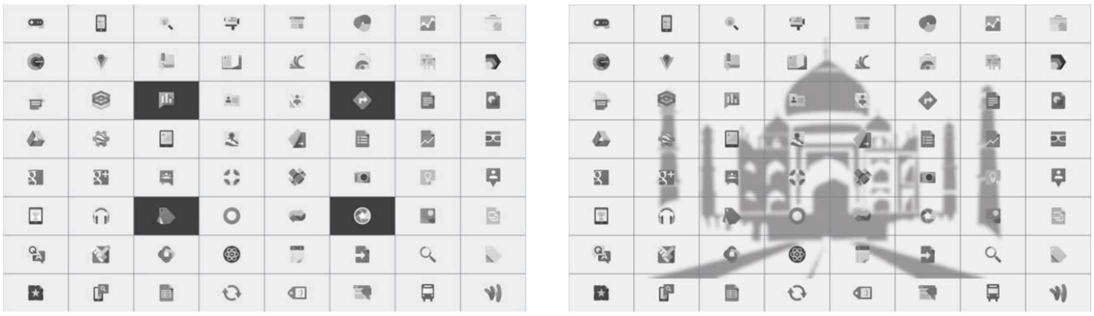


Figure 2.2: Standard menu grid with anchor icons (left) and pictures (right) as artificial landmarks, adapted from Uddin et al.'s [32] experiment

Fine and Minnery [12] investigated how saliency affects one's ability to remember target locations in a map. They found that the ability to recall locations of targets increases with target saliency, and that the effect becomes more apparent as the number of item locations increases. Similarly, Santangelo and Macaluso [29] showed that users memorize locations of items better when they are placed over salient background.

Fujimoto et al. [13,14] investigated how AR label placement affected user memory. They found that users are better at recalling labels placed in close proximity to the target than labels presented in a fixed location (Fig. 2.3).

On the other hand, various techniques can induce an effortful training from users instead of the system adjusting the ease of the process. One of the oldest and most effective methods to enhance one's ability to recall information is the Method of Loci (MoL), also known as the "memory palace" technique [33]. Users memorize information by imagining themselves walking around a familiar environment, and placing items that they need to remember along the path. By walking along the same path and retrieving the items along the way, users can recall what they tried to memorize. While effective, MoL requires many hours of intensive training [21]. Even so, MoL remains one of the most widely used memorization techniques in present times. In an exploratory study by Raz et al. [26], they observed how a person with superior memory was able to recite 540 digits of pi using MoL.

Legge et al. [21] asked users to traverse a path in a desktop virtual environment and to use it for MoL. They found that the virtual environment could be used

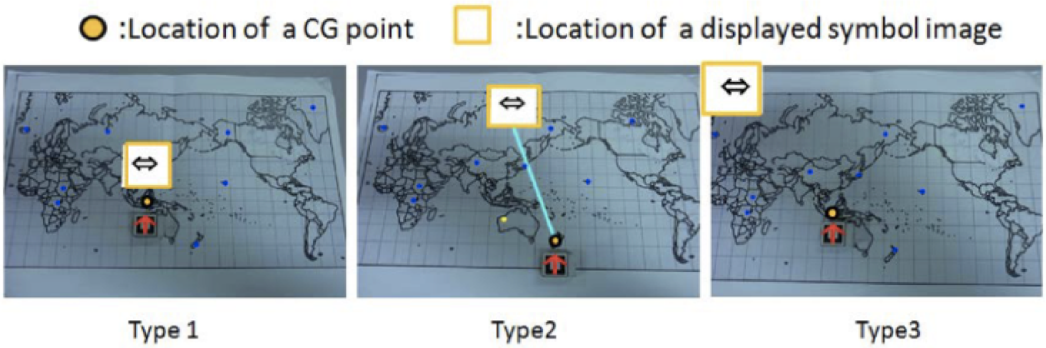


Figure 2.3: Different types of annotation proximities to target locations in Fujimoto et al.'s [13] experiment where the information is displayed: near the location of each CG point on the paper (Type 1), at random locations on the paper (Type 2), and at the same location on the 2D display (Type 3).

as effectively as a familiar environment for MoL. Krokos et al. [19] investigated if immersion had an effect on a virtual MoL setup. They tested 40 participants with the task of memorizing numbers and faces on picture frames placed all over a virtual environment. They found out that using HMDs were better than a desktop display in this scenario.

Rosello et al. [28] created an AR version of MoL, where users could place objects into their surroundings to effectively create a memory palace. Their results show that using AR assisted users in creating a path and applying MoL. As a result, users could memorize items better than with conventional learning.

Cockburn et al. [7] designed a virtual keyboard interface that required users to brush off virtual "frost" in order to see the key label underneath. The frost re-appeared after some time. Their system thus encouraged users to memorize the locations of the keys to avoid having to repeatedly remove the frost, which improved their spatial memory. Ehret [11] showed that spatial memory can also be enhanced through repeated searches for an item.

Based on this observation, Kaufmann and Ahlström [17] investigated whether the use of a projector phone with a peephole interface can yield better spatial memory performance in a map navigation setting, as opposed to using a smartphone with a touchscreen interface. Their results showed that spatial memory performance is indeed better in the projector phone condition. Gacem et al. [15] investigated the impact of a projection-based visual guidance system on spatial memory, with consideration for type of control (user or system controlled) and kinesthetic feedback. They found that the system-controlled setup yields higher recall rates, and that kinesthetic feedback has no effect on recall rates, which contradicts previous findings, e.g., [31].

To summarize, systems which improve users' memorization employ different strategies, whether inherently with the application's display, or externally with the user's interaction coupled with the system's particular device(s).

2.3 Training Environments with Head-Mounted Displays

Augmented and Virtual Reality can assist the training process by presenting guidance to trainees, or simulating a variety of situations that could not be presented in reality. While in some cases, trainees can experience VR and AR applications through smartphones [17] or projector systems [15], they can also wear head-mounted displays (HMDs) for a hands-free experience. Companies like NVIS*, Microsoft†, and Daqri‡ showcase their products highlighting potential for industrial applications.

In manufacturing scenarios, Büttner et al. [5] visualized the design space for AR and VR training, classifying previous studies according to the device used and domains in which they are used (Fig. 2.4). In the analysis, while AR projection systems are more commonly utilized in manufacturing, AR HMDs come in second.

A study by Newman et al. [25] about taxicab driving in VR synthesizes the spatial memory support approaches enumerated earlier. They investigated how altering landmarks (salient objects) and layout information of the virtual town affected how participants navigated their way to drop off virtual passengers. After a few trips, they observed that participants' navigation time decrease as they go along the same route. However, the skill did not transfer upon introducing a new

*http://www.nvisinc.com/
†https://www.microsoft.com/en-us/hololens
‡https://daqri.com/worksense/

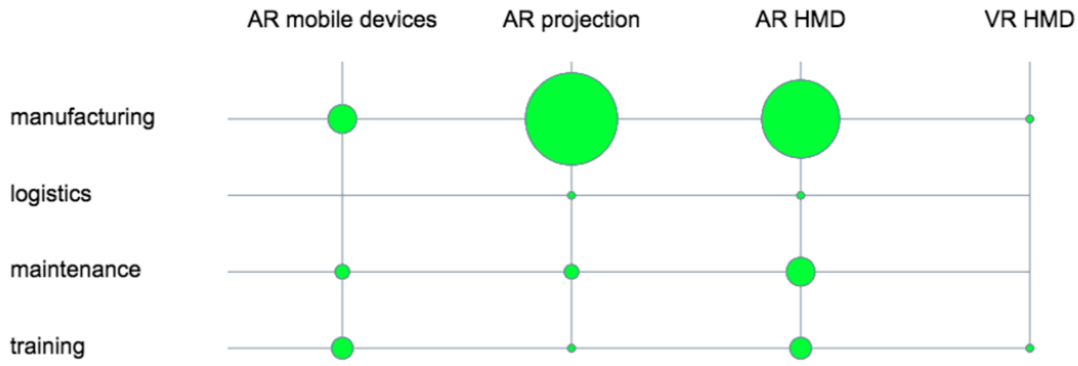


Figure 2.4: Visualization of the design space of AR/VR applications for assistive environments in manufacturing, taken from the analysis by Büttner et al. [5]

town and a new layout, even with very similar landmarks. While offering insight to how humans navigate and memorize virtual spaces, all events happened only in the virtual world.

Bliss et al. [3] compared different training strategies (blueprint, virtual reality, and a no-training control group) in a between-groups study involving a firefighting training scenario. After undergoing training, they asked participants to go through an actual building and checked if there was training transfer. The skills involved in their scenario include memorizing a spatial map of the building. However, the rendering of the building in their VR condition was still rudimentary, focusing only on the building's structure rather than appearance. Similar to the experimental setup problem of Dünser et al. [10], users have different views of their training material. The blueprint version only provided a top view schema of the building, which is different from an immersive experience of a VR application with a perspective view.

Neumann and Majoros [24] focused on manufacturing and maintenance scenarios for AR applications, and identified the different cognitive, performance, and systems issues that came along when using such applications. They have created an application simulation based on transport aircraft maintenance, While they provided illustrations of what their application looks like and how it works, there was no user study to investigate the effectiveness of their setup.

Although there have been studies that explored the benefits of VR and AR in spatial memory and task training, to my knowledge, I am the first to compare the effectiveness of immersive VR and AR for spatial memory training. I aim to answer this question by conducting the experiment described in the next section.

3 Experiment

The goal of this experiment was to compare the positive training transfer of AR and VR training. This section describes the design of the experiment platform, the training procedure, the testing procedure, and the implementation of the training.

To prevent external effects on the training process, like muscle memory from walking to a specific lane in a warehouse, I designed a control room scenario, where participants have to memorize the function of a variety of control buttons and screens. This scenario is very similar to the guidance system used by Gacem et al. [15], where the participants memorized a set of labels and their corresponding highlighted target one by one (Fig. 3.1).

3.1 Experimental Platform

A variety of factors can affect the results of the training, such as the field of view (FoV) and resolution of the display, the realism of the scene, or the training scenario, to name a few. In the experimental platform, I tried to eliminate any external factors that could affect the results.

3.1.1 Room Setup

In actual scenarios, control rooms have varying surface colors and structures, that can affect how well users can read labels that come along with control room elements. Thus, I considered the legibility of the text against the background. DiDonato et al. [9] showed that users read texts better when it is in high contrast against flat surfaces. To help participants see the different targets on the wall and focus on the legible text labels, I printed the layout in gray-scale on white background.

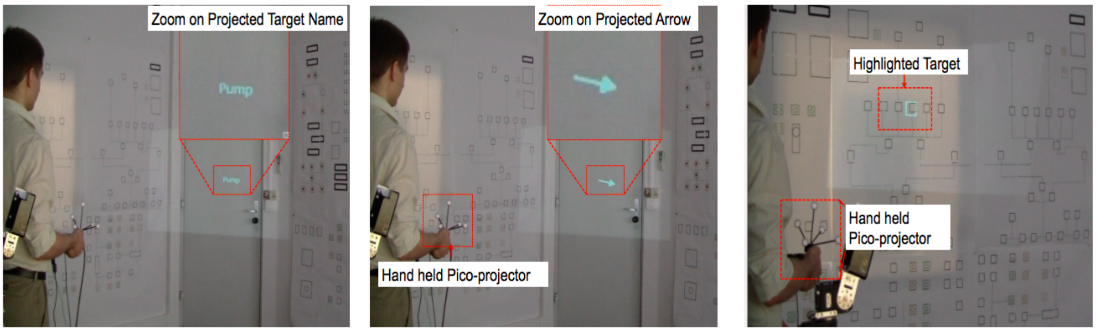


Figure 3.1: The illustration taken from Gacem et al. [15] of the elements of their guidance system, where a handheld projector guides a user to view a direction indicator and corresponding label (left), changes the label from name to a direction indicator in a larger scale as the user moves away from an initialization area (middle), and displaying an outline to the highlight the target element (right).

The actual environment for this study is a control room consisting of 3 walls. The left and right walls measure 1.69m x 4m and the front wall measures 1.69m x 3.78m. On each wall, I placed posters with a variety of targets printed onto them (Fig. 3.3). For the AR condition, the participants should be able to see this control room. For the VR condition, I replicated the appearance of this room as a virtual 3D model (Fig. 3.2).

A generic design of a commercial airline cockpit interface became the basis for each wall of the control room. I enlarged each element of the setup such that even the smallest components were visible from the center of the room. This resulted in a variety of elements, such as buttons, switches, and knobs, that were arranged in grids as well as standing out elements of varying sizes. Overall, the design had 17 unique elements with sizes ranging from 8x8 cm to 10x12 cm. Examples of the various elements are shown in Figures 3.2 and 3.6. Each control element has least 4 cm spacing on all four sides to prevent erroneous memorization due to minute misalignments of the computer graphics (CG). An icon placed at the center of the frontal wall represented a monitor. This area served as an initialization area, where participants could see the label whose position they would learn next.

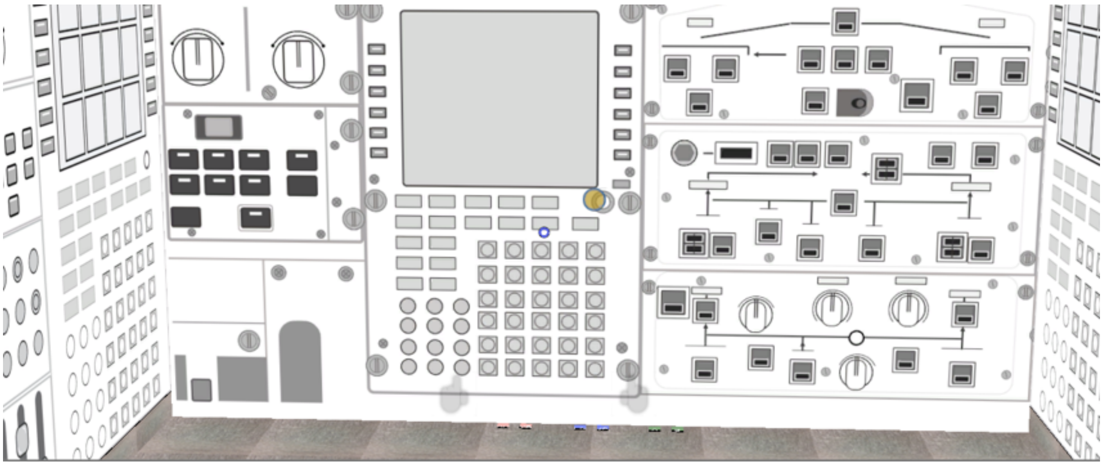


Figure 3.2: View of the control room rendered as a 3D model in the VR condition

3.1.2 Hardware

A variety of factors can affect the quality of AR and VR training, such as the FoV, the resolution of the display, and image quality [20]. To keep these factors constant, participants would use the same HMD for both training conditions. The experiment used a Microsoft HoloLens, an optical see-through head-mounted display (OST-HMD) that does not occlude the surrounding scene. For the VR condition, a black cover blocked the participants' view of the real world environment, and allowed them to see only the virtual room (Fig. 3.4a). To ensure that participants had the same FoV in both conditions, the AR condition used a second cover. This allowed participants to see only the area that would be augmented by the HMD in the AR condition (Fig. 3.4b). In both scenarios, the CG was rendered according to the positional tracking of the device.

3.1.3 Implementation

I developed the experiment system with Unity 2017.1.0.f3 *. During the VR training and short-term memory tests, the training area was a virtual replicate of the scene which was enabled, thus visible to the user. The replicate of the scene was disabled, thus not rendered, during the AR condition to allow a view of the

*https://unity3d.com/

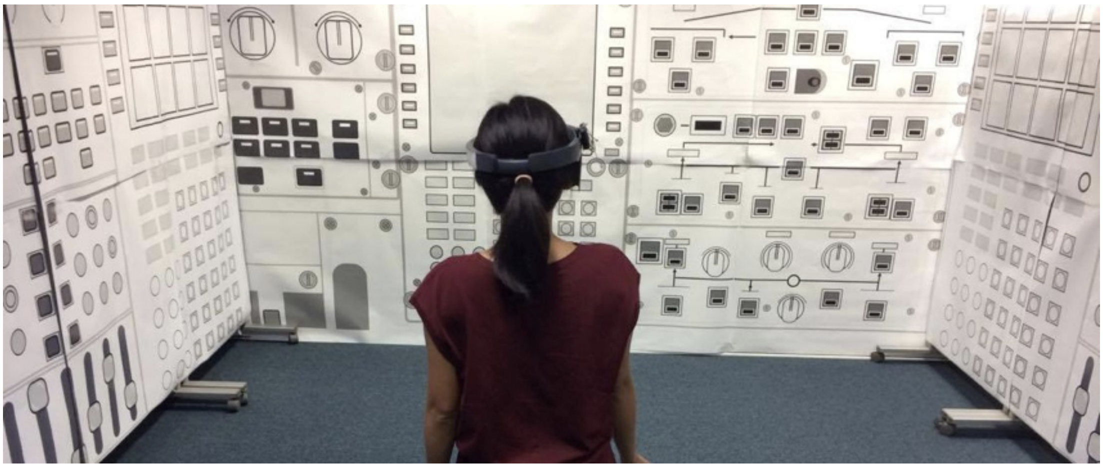


Figure 3.3: View of the control room with paper printouts on walls in the AR condition

training room with CG overlaid onto it.

3.2 Subjects

For this study, I recruited 16 participants from the Nara Institute of Science and Technology (13 male, 3 female, age 22 to 33 (25.75 ± 3.044)). All participants received compensation for their time with an amount equivalent to 10USD. Before taking part in the experiment, participants signed a consent form and answered a preliminary survey about their familiarity with AR, VR, as well as the Microsoft HoloLens. Results show that 9 participants (56.25%) had prior experience with AR, 10 participants (62.5%) had prior experience with VR, and 5 participants (32.25%) have used a Microsoft HoloLens in the past.

3.3 Procedure

The participant followed certain steps to complete the experiment. This section elaborates these steps with an overview of the experiment procedure, and a detailed explanation of each component.

(a) VR (b) AR

Figure 3.4: Microsoft HoloLens masks for the two conditions, with the area augmented by the CG (inside the yellow box) remaining open in AR, and closed (full peripheral FoV occluded) for VR.

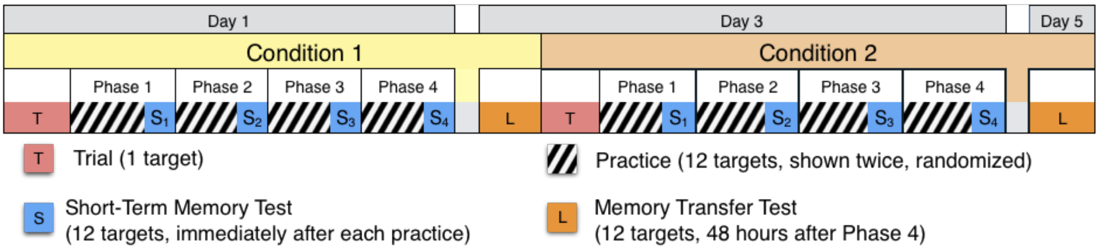


Figure 3.5: Experiment procedure mapping the two sessions over the course of five days.

3.3.1 Overview

The goal of the study was to evaluate the quality of memory transfer. I conducted a within-subjects experiment over the course of 5 days. Fig. 3.5 illustrates the structure of the experiment timeline. To know if they had any experience with AR or VR, participants completed a survey before starting the experiment. On the first day, participants would complete a trial session, denoted by T in Fig. 3.5, that consisted of learning the position of a single label to familiarize themselves with the system. They then took part in 4 training phases, with each phase consisting of a practice session followed by a short-term memory test (denoted by S_i in the figure, with i being the phase number, e.g., the fourth short-term memory test in the sequence is denoted by S_4).

On the third day, participants came back to take a memory transfer test, denoted by L in Fig. 3.5. After they completed the test, they started the second training session in the untrained condition. Finally, on the fifth day, participants came back for a memory transfer test of the second condition and provided feedback via free-form comments.

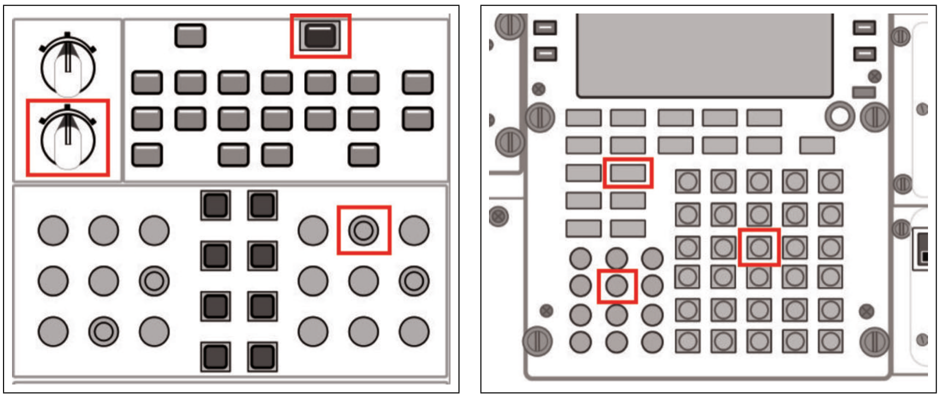
As per instruction, participants stood in the center of the room so they could see all walls during the training. However, they were not prevented from taking some steps to the side, as this would be natural behavior that is possible in both scenarios. Furthermore, the experiment did not require participants to perform any motion for kinesthetic memorization.

In the control room layout, I selected 24 evenly distributed targets on the walls, chosen in terms of location and saliency. Hereby, saliency is defined by the uniqueness of an object (see Fig. 3.6). For example, an object that is part of a grid pattern is expected to be less salient than an object that is standing out distinctively. Similarly, a smaller object is less salient than a similar larger object. I assigned each target with a unique label based on terms related to aviation. I sorted these targets into two datasets, wherein participants learned one dataset in the VR, and the other in the AR condition. I also classified these targets into near and far targets, according to proximity from the initialization area.

During each training session participants would see each label twice and the order in which the labels appeared was randomized. During the testing phase the application randomized all 12 targets covered in the training session and tested the participant. For each part of the experiment, I also recorded the time participants took to complete the task, the number of correctly answered questions, as well as the errors.

3.3.2 Training Procedure

In this study, the learning procedure consisted of three steps. For each step, there is a corresponding time component of the training session adopted from the experiment by [15]. It should be noted that Fig. 3.7 to 3.9 provided came from Microsoft HoloLens' live capture of the training application.



(a) Salient (b) Less Salient

Figure 3.6: Examples of salient and less salient targets.

1. The target icon appears in the center of the initialization area of the front wall (Fig. 3.7). The target icon stays in this position until 5 seconds have elapsed or the user has already clicked a button on a controller. The time elapsed since the target icon appears is called the stimulus time.

2. The target moves from the central area towards the goal location (Fig. 3.8). Whenever the target would leave the user's view, it would stop. The total time it took for the target icon to move from the initialization area to the position of the destination target is called the movement time.

3. After reaching its destination target, the icon changes color and disappears after 5 seconds, or after the user clicked a button on a controller (Fig. 3.9). The time it takes for the target icon to disappear from the destination is called the highlight time.

The target icon consisted of a label and an indicator that would show the direction towards the corresponding element on the wall (Fig. 3.7). The label and the indicator have colors to improve legibility and help participants see them against the wall. In an ideal scenario, with a wide field-of-view HMD it would be possible to simply present the labels at the corresponding location and present only minimal guidance to help users build a spatial map.

However, due to the limited field-of-view of the device used in the experiment

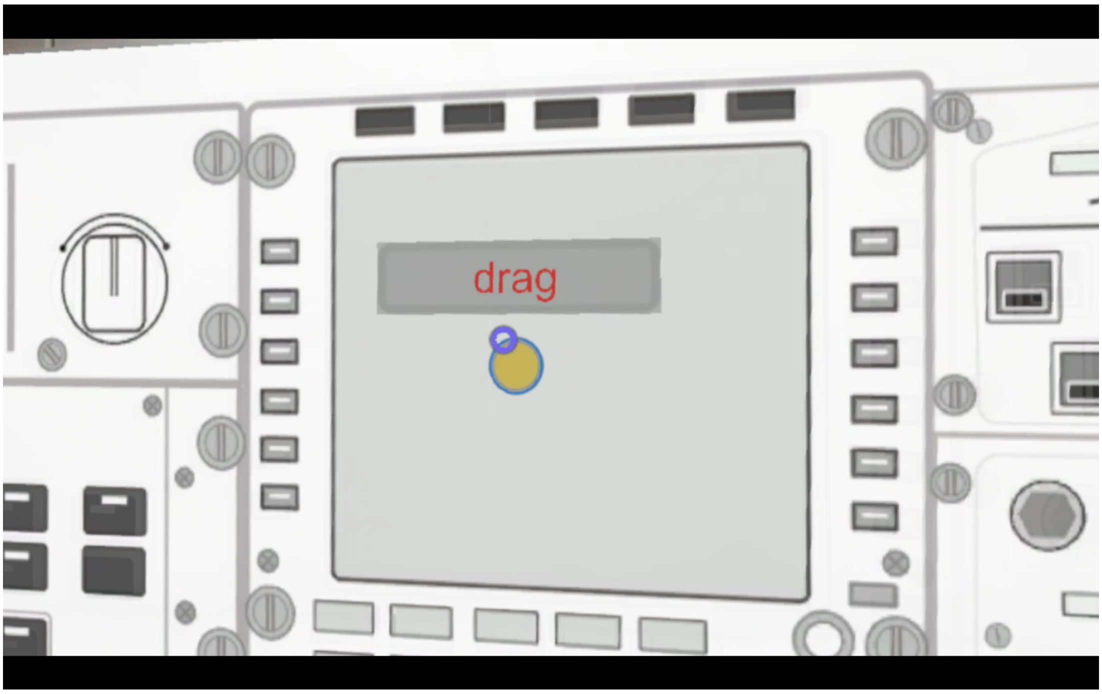


Figure 3.7: The label and the icon in the source position.

the system presented the labels at an initial location. Participants followed each label towards the corresponding location, but contrary to [15], the indicator and the label moved together to the target location. The basis for this was that users memorized items better, if the label was close by [13].

3.3.3 Testing Procedure

To verify the results of the training, the short-term memory and a memory transfer test evaluated how many targets the users have memorized. During these tests, participants had to point out the location of the elements that corresponded to labels they had learned during the training. The participants answered the short-term memory test in AR/VR. Here participants had to align an icon with the corresponding element. After they confirmed their selection, the icon changed to show them whether their selection was correct (Fig. 3.11).

On the other hand, the participants took the training transfer test in the physical environment (regardless of the training condition). For each label, the tester

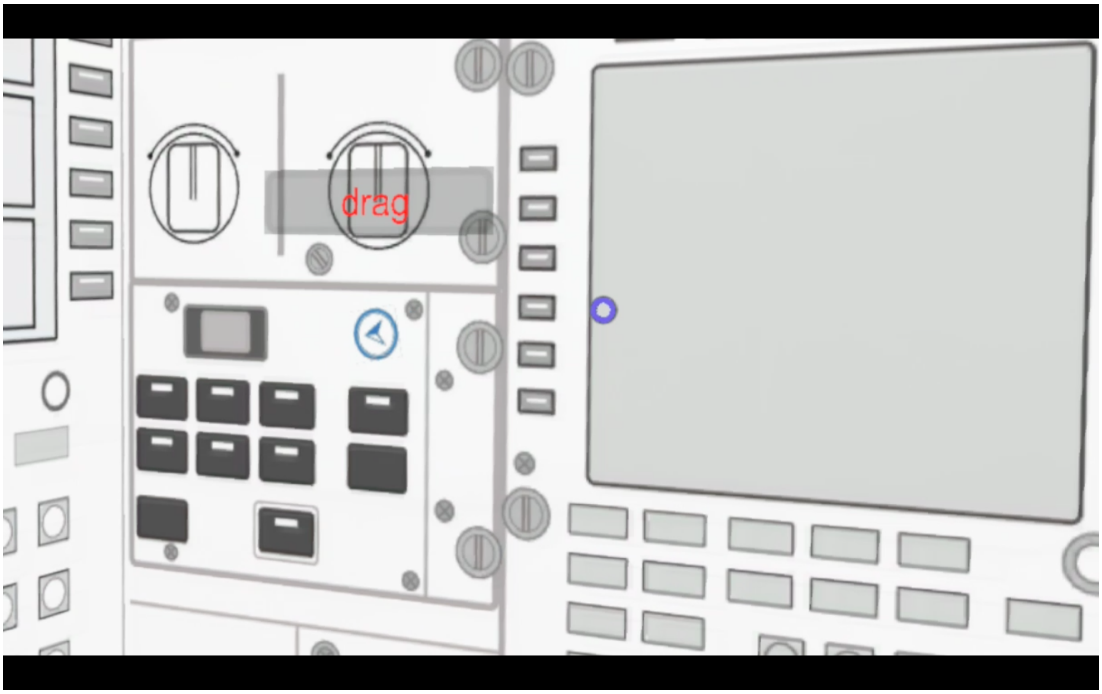


Figure 3.8: The icon changing from a position to an direction indicator.

asked participants to point out its location on the wall and provided no feedback about the correctness of the result. Therefore, participants would not rely on the familiar training device, environment, and process of elimination to apply the learned skill.

The experiment conducted by [15] became the basis for the number of targets per practice, as well as frequency and timing of the tests. These choices were also consistent with findings by [8], where a thorough review of related literature synthesized the definitions of short-term memory and long-term memory in terms of temporal decay (duration) and ability to store chunks of information (capacity). Phases 1 to 4 containing the practices and short-term memory tests have no breaks, so as to remove effects of interference. This makes 48 hours an adequate gap between the end of Phase 4 and the start of the memory transfer test. Finally, I chose 12 targets, which is more than the average 7 chunks that people can accommodate. The expectation is that participants are unlikely to achieve a 100% recall rate on S_1.

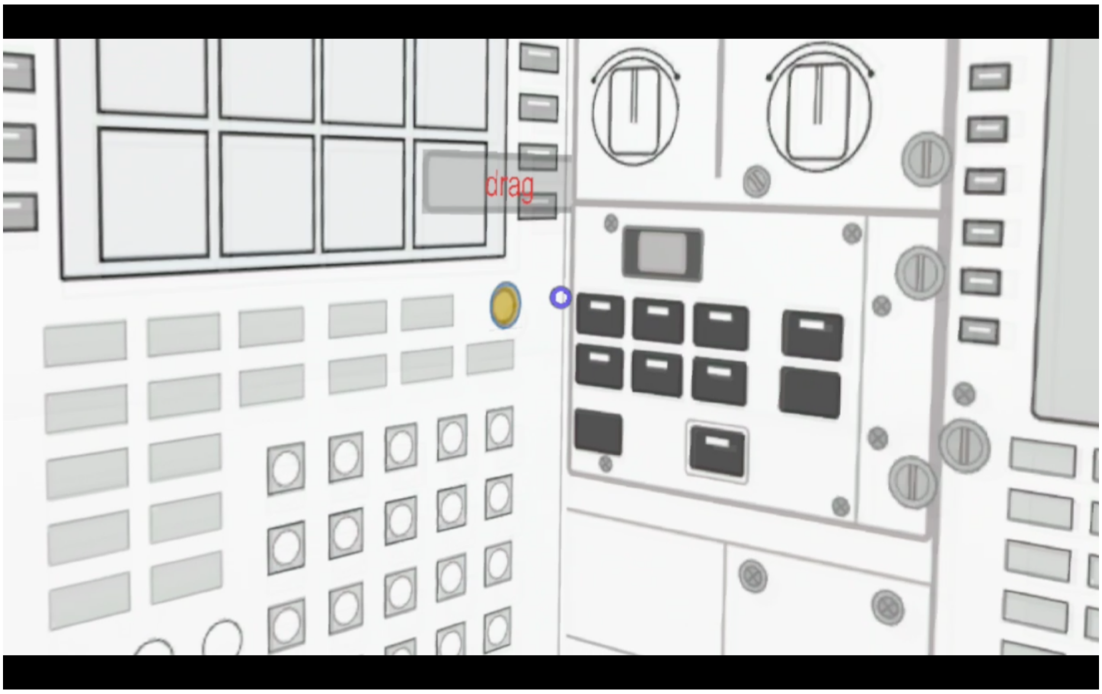


Figure 3.9: The label and location must be memorized within 5 seconds.

3.4 Independent Variables

The experiment was counterbalanced in terms of the order in which participants trained each condition (AR or VR first), as well as what data set was used to ensure that this did not affect the results. This study had the following independent variables:

Condition: Denotes whether participants were training in AR or VR.

Proximity: Denotes how close the target was with respect to the initialization area. I define the target to be near to the starting area if it is on the front wall or is no more than 1.33m away from it (corresponding to the third of the side walls that is closest to the front wall). I consider the rest of the side walls as far regions. Fig. 3.12 shows the areas of near and far targets. For each target, this proximity is the same all throughout the training sessions, as the positions of the initialization area and each individual target are fixed. This should not be confused with the proximity of annotations in [13], which focused instead on the distance between the label and the target.

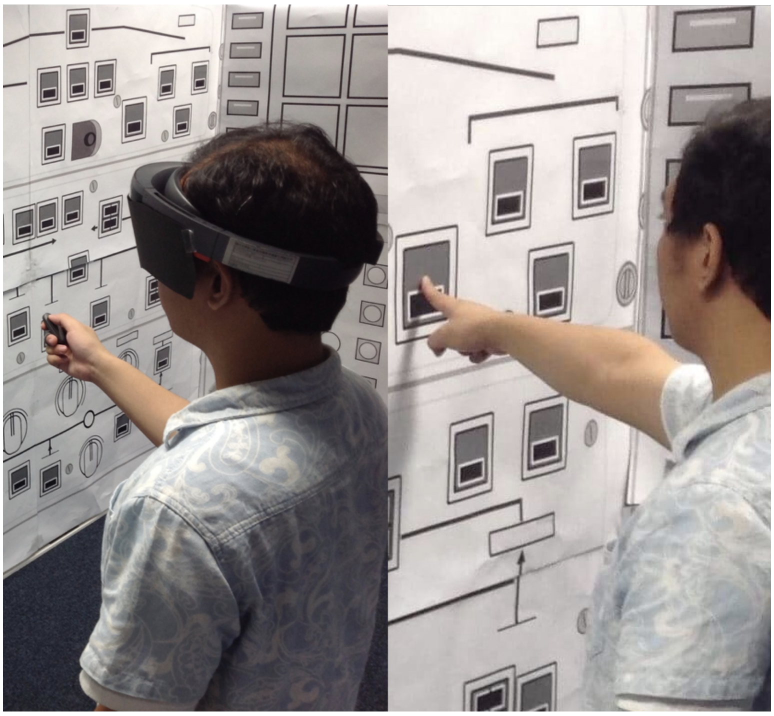


Figure 3.10: Participant with the device during short-term test (left) and without the device during the memory transfer test (right) .

Saliency: Denotes if a target was unique or large and thus *salient*. On the other hand, targets that were part of a grid and smaller were *less salient*(Fig. 3.6).

Session: Denotes when the test scores were collected. In particular, these are the four training tests S_1, S_2, S_3, and S_4, and the memory transfer test L.

3.5 Hypotheses

While the experiment attempts to investigate the difference between VR and AR conditions, I designed the training scenario to be as consistent and controlled as possible for both environments. I also tried to match the 3D model of the room with the actual room in terms of layout and positions of all elements. With this in mind, I expected participants to retain information equally well during the training for AR and VR (**H1**).

Another expectation is that training in the actual environment will facilitate

(a) Correct (Green) (b) Wrong (Red)

Figure 3.11: System feedback after the user selects a location in the short-term memory test.

memory transfer. The AR condition lets the participants view the actual environment during training, contrary to the 3D rendering of the walls occluding the actual environment in the VR condition. It would make sense that training transfer is easier in AR because the extent of the world knowledge is the same (i.e., the viewed environment during training is the same as when the actual task is performed). Thus I expected participants to recall labels they learned in AR better than those learned in VR during the long-term test (**H2**).

Furthermore, based on previous findings [13], I expected that the distance from the initialization location to the target location will not affect how well participants memorize the location because the labels travel along with the icon towards the target location (**H3**).

Finally, I expected the saliency of the targets to affect how well the participants memorize them. (**H4**). Based on the definition of saliency provided earlier, targets with a bigger size would catch more attention, thus, participants could pick up the location information much faster. The elements designed to stand out over a structured grid is also expected to emulate the artificial landmark observation by [32], which should yield better recall rates for these targets.

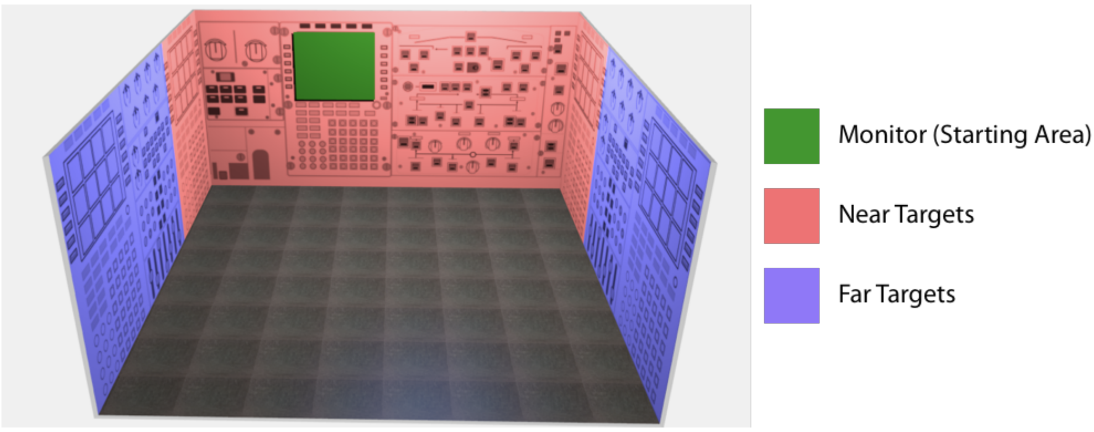


Figure 3.12: Bird's eye view of the control room scenario showing target areas divided by proximity.

4 Results and Discussion

The design of this experiment involved a within-subjects study. This section explains the corresponding methods I have used in processing the data I gathered from test scores and training times. Afterwards, the discussion portion addresses the original hypotheses, as well as related points and observations that can explain the observed results.

4.1 Experiment Results

I recorded the number of correct responses of the participants on all tests (S_1, S_2, S_3, and S_4 and L) on both conditions (AR, VR). The recall rates for each test are shown in Fig. 4.1.

The analysis of training data used repeated measures ANOVA with *Condition* × *Saliency* × *Proximity* × *Session*, using all 4 training sessions. The results show that *Session* ($F_{3,45} = 102.320, p < 0.0005$) and *Saliency* ($F_{1,15} = 15.882, p < 0.001$) had a statistically significant impact on the results during the short-term training tests. I found no statistically significant interactions between the different factors.

For the comparison of the performance between the individual tests, one-way ANOVA test revealed a statistically significant difference between S_1 and S_2 ($F_{1,15} = 96.71, p < 0.001$), S_2 and S_3 ($F_{1,15} = 43.55, p < 0.001$), and S_4 and the memory transfer test L ($F_{1,15} = 22.16, p < 0.001$).

To better understand the results, one-way ANOVA tests showed how *condition* affected the results at each training stage for each session. It also revealed a statistically significant difference between the AR and VR conditions in the scores for S_4 ($F_{1,15} = 4.88, p = 0.04$). However, it did not show any statistically significant difference between participants' ability to recall *salient* targets compared

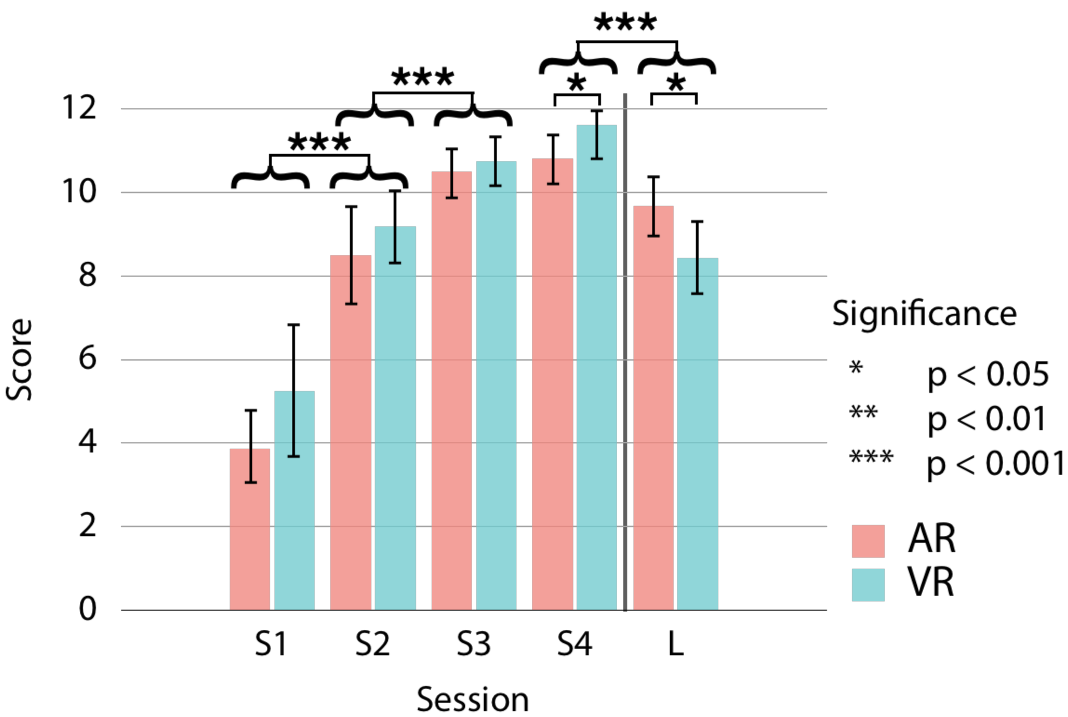


Figure 4.1: Mean scores per condition across all short-term tests (S_1-S_4) and the memory transfer test (L). The whiskers denote confidence interval.

to *less salient* targets for S_4.

The evaluation of the results of the memory transfer test employed a repeated measures ANOVA *Condition* × *Saliency* × *Proximity*. The results show that *Condition* had a statistically significant effect on the results ($F_{1,15} = 6.229, p = 0.025$). I also found a statistically significant interaction between *Saliency*, *Condition*, and *Proximity* ($F_{1,15} = 7.3, p = 0.016$). A post-hoc Tukey Test showed that there was a statistically significant difference between users' recollection of far salient targets that were trained in AR and near less salient targets that were trained in AR.

I also checked whether prior experience with the device had any effect on participants' performance, regardless of AR or VR condition. This study used a t-test to compare scores between participants with and without prior experience. I used the mean scores for short-term and memory transfer tests for this. Although the sample size of the experiment is too small to be generalizable, the results reveal no significant difference on both the last short-term memory test S_4 (VR: $t = 0.13, p = 0.90$; AR: $t = -0.75, p = 0.47$) and L (VR: $t = 0.04, p = 0.97$; AR: $t = -0.82, p = 0.43$).

This experiment used two sets of target locations. I compared the two datasets in terms of difficulty with one-way ANOVA, and I found no statistical difference between the results of the two datasets ($F_{1,15} = 2.00, p = 0.18$). The order in which participants trained in the conditions also did not have a statistically significant impact on the results ($F_{1,15} = 1.15, p = 0.3$).

Table 4.1 shows how the *proximity* of the targets to the initialization area affected their recall rate. Repeated measures ANOVA did not reveal any statistically significant interaction between *session* and *proximity*.

Finally, we also investigated how the *saliency* of the targets affected participants' recall rates (Table 4.2). Repeated measures ANOVA did not reveal any statistically significant interaction between *saliency* and *condition*.

4.2 Training Duration Results

Both VR and AR sessions had similar training and testing durations. On average, training and testing lasted for 21.94 minutes, and the memory transfer test took 4.06 minutes in the VR condition. Meanwhile, in AR the training sessions lasted for 19.31 minutes, and the memory transfer tests for 4.75 minutes. A paired t-test reveals that there is no significant difference between the two conditions (Training: $t = -1.324, p = 0.205$; Memory transfer test: $t = 0.983, p = 0.34$).

The training application also logged time durations (in seconds) during the training sessions. Adopting a similar approach made by Gacem et al. [15], the

Table 4.1: Average recall percentages according to proximity.

Proximity	**Short Term**				**Memory Transfer**
	S1	*S2*	*S3*	*S4*	
Near	36.89%	76.89%	88.00%	94.22%	73.78%
Far	38.36%	69.81%	89.94%	93.08%	77.36%

training duration (i.e., global time) of memorizing one target has been divided into three: stimulus time, movement time, and highlight time. For the purposes of this study, the global time is the total of these three durations in one session. The graph shows how these three durations combine (Fig. 4.2).

Stimulus time is the duration in which the name of the target is displayed until five seconds have elapsed. Movement time, on the other hand, is the duration in which the application show the target as it approaches the target location. Finally, highlight time is the amount of time spent by participants looking at the highlighted target before starting the next trial, which should be no more than 5 seconds.

I used repeated measures ANOVA for all duration divisions (stimulus, movement, highlight) and global times. For stimulus time, results showed a statistically significant difference for *Condition* ($F_{1,15} = 8.398, p = 0.00447$) but not for *Session*. On the other hand, the difference in movement time was not statistically significant enough for *Session* ($F_{3,45} = 2.586, p < 0.0563$). Post-hoc Tukey tests revealed that movement for S_2 was slower than S_4.

In terms of highlight time, a statistically significant difference exists for *Session* ($F_{3,45} = 31.022, p < 0.001$). Finally, for global time, a statistically significant difference exists ($F_{3,45} = 13.208, p < 0.001$) for *Session*, but not for *Condition*, however noting that $p = 0.0719$.

4.3 Discussion

In a comprehensive review, Krujiff et al. [20] enumerated issues that relate to perception in AR. Afterwards, the authors classified the issues into categories

Table 4.2: Average recall percentages according to saliency of target.

Saliency	**Short Term**				**Memory Transfer**
	S1	*S2*	*S3*	*S4*	
Salient	43.17%	77.97%	92.07%	94.71%	79.30%
Less Salient	29.94%	67.52%	83.44%	91.72%	70.70%

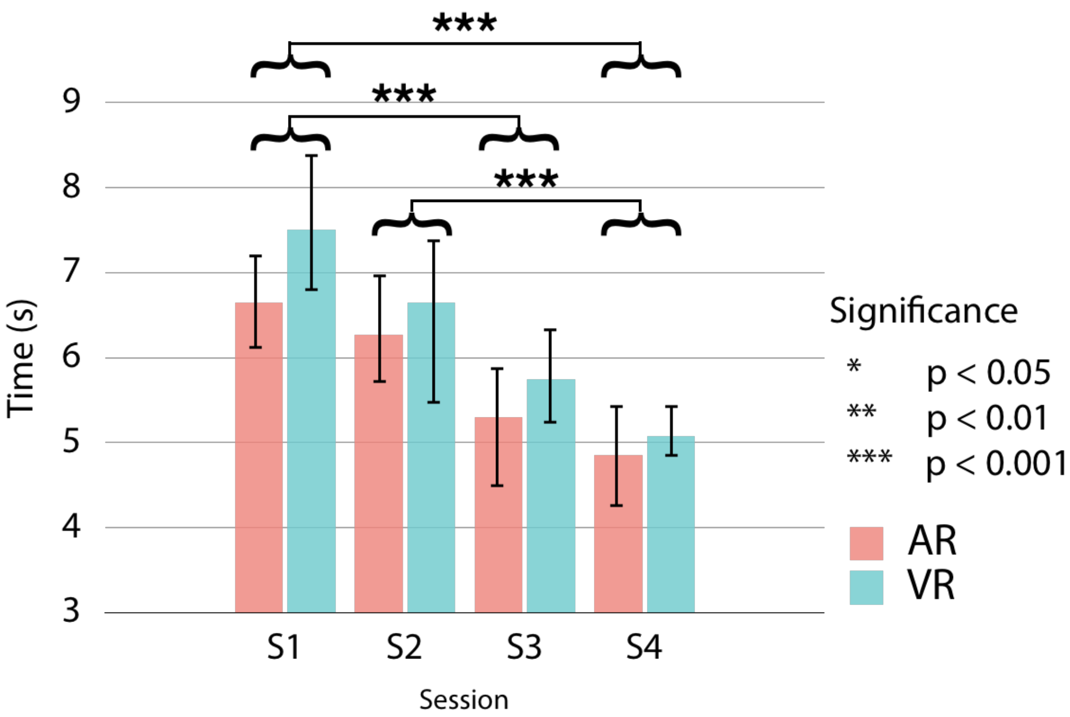


Figure 4.2: Average times per practice session (global time) with corresponding whiskers to denote confidence interval.

involving environment, augmentation, display device, and the user. Even if their review focuses on AR applications, some of the issues overlap with the perceptual issues in VR conditions. The relevance of these issues are important in discussing the results of the experiment. The results show different interpretations for short-term test scores and memory transfer test scores. In this section, I discuss how training duration results can provide insights on the participants' performance, how the results supported or rejected the hypotheses, how the feedback from participants revealed different approaches to the memorization task, and how memory lapses affected the scores by analyzing participants' errors.

4.3.1 Training Duration Analysis

The results obtained from training duration logs show varying behavior for each segment of the training phases. As *Condition* affected stimulus time, participants

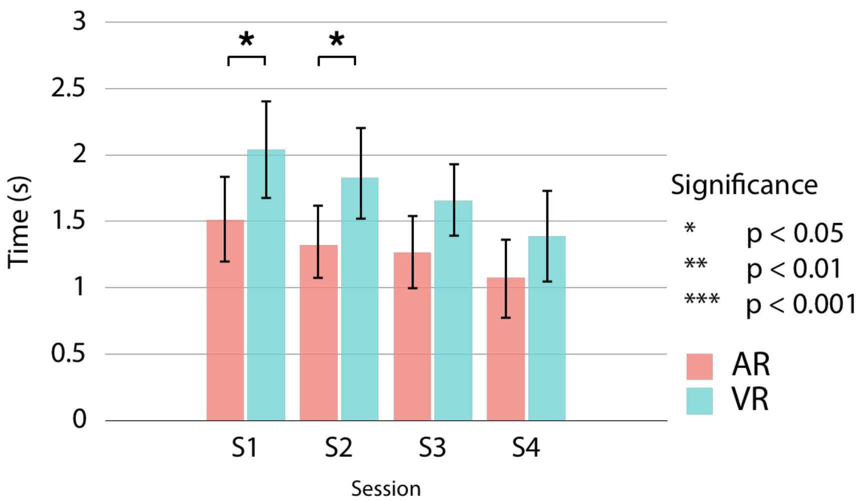


Figure 4.3: Average time durations per practice session when the target still has not moved from the initialization area (stimulus time) with corresponding whiskers to denote confidence interval.

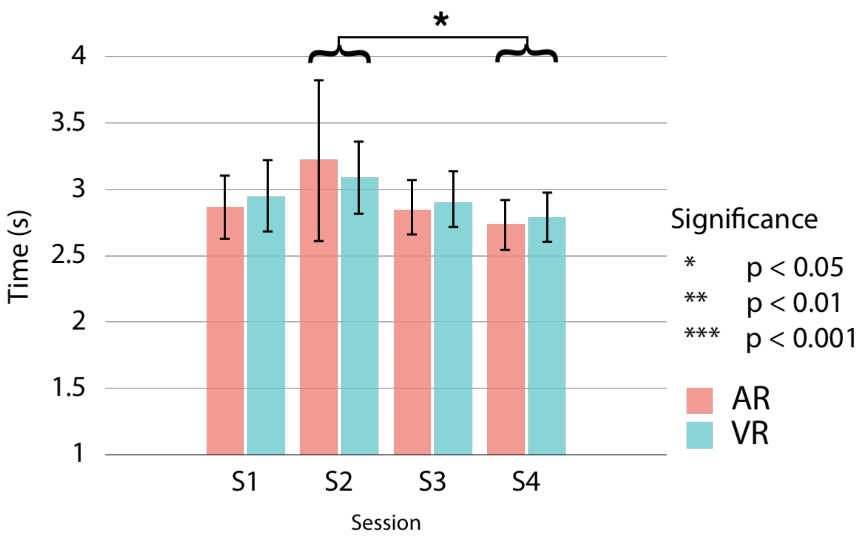


Figure 4.4: Average times per practice session when the targets are moving from initialization area to respective target locations (movement time) with corresponding whiskers to denote confidence interval.

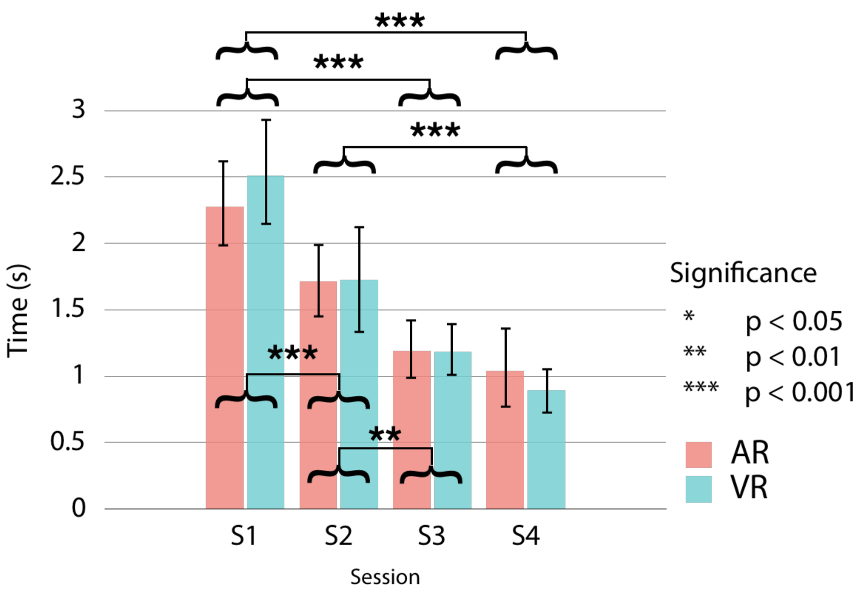


Figure 4.5: Average times per practice session when the labels stop at target locations (highlight time) with corresponding whiskers to denote confidence interval.

started to proceed with the navigation more immediately in the AR condition. Having no statistically significant difference among movement times, the participants are more likely to follow a target as it moves rather than stopping. Interestingly, the highlight time results revealed that participants were more likely to proceed immediately with the next target or end the session, as they may have already finished memorizing the label-target association early on. Finally, even if global times between AR and VR do not have a statistically significant difference, the graph (Fig. 4.2) illustrates that on the first take, participants took more time in the VR condition. Furthermore, the global time decreases as each training session passes, effectively making the average duration of the last training session under 5 seconds, for both conditions.

4.3.2 Hypotheses

The results of the last short-term test reject hypothesis **H1**. The graph from Figure 4.1 shows that VR training consistently outperformed in the short-term tests. Even if the results did not show enough statistically significant difference between VR and AR in average global times during practice sessions ($p = 0.0719$), the global times for the VR condition were longer than the AR condition. This means that users spent more time in VR than in AR. While this may be a potential indication of more effortful training in VR, the case of slower task times can also be attributed to a more difficult time distinguishing the CG augmentations with the 3D model version of the room, which is also a CG augmentation.

However, the effort that participants had with the rendering similarity could be eased because the visual information only presented the rendered CG without lighting and shading factors that are present in the AR condition. This simplified training environment may have also contributed to less distractions, and better performance in the short-term tests.

On the other hand, participants showed a sharp drop in the recall rate when moving from the VR to the real environment during the memory transfer test. Although the performance for the AR condition decreased as well, the recall rate was significantly higher than for the VR condition. This supports hypothesis **H2**. Most likely, the association of the training and the testing environment seems to have been stronger in the AR condition, which led to better memory transfer. This implies that training in VR must be complemented with practical sessions to assist the memory transfer. In the AR condition, the CG augmentations may stand out in contrast to parts of the real environment affected by lighting and shadows. This coupled with the fact that participants finished going through the practice sessions faster in AR than in VR, the participants may have found the contrast as an easier navigation mechanism to follow a label and not lose sight of it during training.

As mentioned in the experimental design earlier, it is also important to analyze the design of the training environment, aside from just comparing the scores in terms of *condition*. Although no statistically significant effect of the training condition exists on the recall rate of salient and less salient features during the memory transfer test (Table 4.2), plotting the corresponding recall rates showed that while salient features were recalled equally well for both conditions, participants recalled less salient features less frequently after the VR training. In the future, it is necessary to conduct further experiments with a larger number of participants, and more complicated tasks to determine if there is a potentially significant effect.

These results contradict the findings of Gacem et al. [15]. This is likely because in their study the label always remained on a single position, while in this study the label moved towards the location. This helped users associate the label with the location [13]. However, even if the results support hypothesis **H3**, there might still be subtle differences among participants when memorizing near or far targets.

Finally, the results reject hypothesis **H4**. I attribute this to how this experiment defined what is salient and what is less salient. Other properties of a target, like the saliency of its outline against the white background, could have assisted the memorization. Complexity of the structured grid design can also be looked at (e.g., number of elements, variations of elements in the structure). While I expected participants will be more likely to make mistakes on targets located in a grid, some participants reported that they memorized these locations like an entry of a matrix (i.e., putting the label, row number, and column number in one chunk). Thus, these participants gave much more effort and attention in memorizing less salient targets which belong to a structured grid. It is also possible that the targets deemed as small were large enough to be easily memorized by the participants.

This study argues that the test results can be explained by how well the training matched the application scenario. While the environment model used in the VR training matched the real scene, it was rendered with simple shading and no illumination. During the training phases in VR, the participants exerted more effort because of the appearance consistency of both the virtual rendering of the room and the label rendering. On the other hand, the environment that the user was presented with in the AR training scenario included shadows, flickering light, and potentially transparent computer graphics. Therefore, it better matched the transfer environment, whilst decreasing the users' cognitive load during training.

As expected, the participants learned the label locations over the course of multiple training sessions, as shown by significant differences between consecutive short-term test scores. The scores plateaued at the end, resulting in no statistically significant difference between the third and last short-term test scores.

However, the differences between the last short-term test scores and the memory transfer test scores revealed different discussion points for the hypotheses.

4.3.3 Observations

In the experiment, I ensured that the virtual room model is a close reconstruction of the actual room. This is supported by some participants mentioning that they did not notice the difference between the two conditions. At the same time, other participants noted that the obvious materials used in creating the actual environment (e.g. printed paper corners, tape) somehow served as landmarks [32]. These material details do not appear in the VR condition, as the 3D model of the wall occluded these features.

Since the level of realism may vary from one training scenario to another, designers must exercise caution when dealing with the issue of realism of the virtual objects and the virtual world. These observations further raise the question of how realism and truthful reconstruction of the scene affect the results as well. Given the results of this study, one would expect that an imperfect reconstruction would lead to even lower retention rates. However, this remains to be verified in the future.

In both conditions, some participants moved one or two steps back in order to increase the area they could see through the display. This suggests that when training on HMDs, the field of view could have an impact on how well users memorize the location of targets, and this investigation is the subject of future experiments.

4.3.4 Error Analysis

As participants made fewer and fewer errors as the training progressed this observation focuses on what kind of errors were made throughout the training. Out of 1,920 trials across all tests (short-term and memory transfer), there were 512 errors. For the first test, the majority of the errors was incorrect assignment of labels to targets of the same data set. For the second data set, participants selected targets that were part of the data set they had previously learned. This is in line with the observations made by Leifert [22], who reported that users are better at recognizing locations than the content associated with them.

Adjacency of targets also had an effect on memorization. In the setup I assigned two big targets that were next to each other and belonged to the same grid, to different data sets. One participant switched the answers and two participants mentioned that they paid special attention to these targets because of the adjacency.

4.3.5 Recommendations and Guidelines

The results reveal differences between the VR and AR conditions in terms of how the participants experienced them. I would like to point out important design considerations that may assist the design of future training scenarios in both conditions. In this section, I would like to offer recommendations about training environment design and training time.

The contrast of the CG augmentations against the environment background can contribute to the amount of effortful training. In the VR condition, since both the CG augmentations and the environment are both virtual, this can easily add a few more seconds in distinguishing where the label is during stimulus time. Even with putting colors on the label and the indicators, I recommend adding subtle cues (e.g., blinking cues surrounding the label, audio) for the users to redirect the attention to the labels. However, if this slower stimulus time is necessary for an effortful training, there should be a mechanism to decrease the allowable time that the label stays on the initialization area for the next rounds of training. I also recommend including a dynamic difficulty adjustment in the training (e.g. faster stimulus time, faster movement time), as different participants have different pacing styles and motivations while training.

On the other hand, since the AR condition already induces a faster stimulus time for participants, the contrast of the room lighting against the CG augmentation might have already been enough. However, I recommend checking first if the lighting conditions in the training area is enough to produce this contrast. In the case of this experiment, the lighting was high enough in the ceiling to still make the printout elements visible while brightening the overall surrounding.

The frequency of the training can also be adjusted, as informed by the results of the short-term and memory transfer tests. Since VR outperforms AR in the short-term memory tests, I recommend that short-term training in VR should be done as a supplement to the AR training. It should be noted however, that this adds to the overall initial training time. This should be done if there is enough time to do VR training, and if the participants are physically comfortable in wearing the HMD for a longer time.

In cases of a short training period, it is only logical to hold the training in the AR condition if the actual environment is also available. This online training would provide better transfer, as the view of the training environment and the actual environment is similar. In case trainees still forget some items or procedures, they can always just wear the OST-HMDs and perform training in the AR condition to refresh their memory.

5 Conclusion

This study compared how spatial memory from AR and VR training transfers to actual applications. Although the results showed that VR performs better in the immediate post-training tests, it was performing significantly worse in the memory transfer test. This suggests that AR is better suitable for training memorization of spatial maps, especially if users cannot perform multiple training sessions. Nonetheless, VR remains a viable tool for improving spatial memory, as it is easier to repeat the training in VR, compared to AR.

With the spread of affordable head-mounted displays on-the-job training of users through AR will become even easier and more appealing, especially if it does not interfere with the overall work flow. It is also preferred to training in VR as it helps retain the learned information and transfer it to the actual scenario. Nonetheless, VR can compensate the training to further reduce the overall training and memorization time.

Additionally, the insights from the the training time durations indicate how VR and AR applications could be differently perceived. Training environment designers and developers who wish to use AR or VR for improving memorization skills should look at how participants exert more effort or take more time in registering and storing information. They should also be more conscious on the learning strategies, as trainees might prioritize spatial information over content when presented with a task that requires memorizing both.

5.1 Limitations

There are several limitations to this study, which will be subjects of future explorations. The nature of the training environments, the display device used, and other considerations limit the generalizability of the study.

5.1.1 Training Environment Design

The analysis of perceptual issues in AR [20] pointed out that augmentations of the real world can disturb the way they can be recorded or perceived. Thus, it is important to point out issues of environment conditions and structure. They find that lighting of indoor environments like this experiment can have a ripple effect in terms of incorrect display of color and incorrect augmentation. While the VR condition only dealt with uniform lighting (i.e., no complex illumination) for the 3D models, the real environment had varying lighting conditions which may affect the user's perception of the label and the wall. There were no surface distortion differences between AR and VR.

In terms of rendering quality, the 3D model generated in the VR condition was not photorealistic. Theoretically, a perfect VR training environment and experience should be identical to an actual environment. In this ideal VR scenario, the experience should be visually realistic. However, in this setup, I have eliminated realistic components in the VR condition. In future experiments, this can be addressed by replacing the 3D model of the VR condition with a camera capture (preferrably 360°) of the actual environment to approximate the lighting and rendering conditions.

According to post-test subjective feedback, a few participants could not easily memorize labels in English. While Neumann and Majoros [24] argued that the design should be adjustable to accommodate different languages for multicultural crews, I would like to design future experiments with uniform cues to have a balanced basis for comparison and observation. This can be addressed by either asking first if the participant is familiar with the labels used in the experiment, or change the labels in the scenario from words to universal symbols (e.g., warning symbols).

5.1.2 Device Use and Capabilities

Kruijff et al. [20] also added another dimension in their analysis: which of the platforms supporting AR are these issues applicable to? They suggested that display types has a major influence on perceptual issues, along with a table that shows the advantages and disadvantages of platforms like head-worn displays, handheld devices, and projector-camera systems. In this experiment, I used a Microsoft HoloLens, an OST-HMD. The advantages and disadvantages that they noted touch upon the same issues I encountered with this experiment (e.g., FoV, display resolution, image quality), as well as other factors that should be taken into consideration in the future.

One thing is the capability of OST-HMDs for "real-world verification". This sets it apart from video see-through head-mounted displays (VST-HMDs) that may add latency, and reduce the feeling of presence in a real environment as users know it is a video capture instead of seeing the real one. Hence, I used an OST-HMD for this experiment, as the video capture of VST-HMDs does not match the real environment shown during the memory transfer scenario. Furthermore, to demonstrate that the VR and AR conditions are different from each other, this real-world verification should be eliminated in the VR condition. As Azuma [1] required the combination of the virtual and real, I removed the view of the real world for the VR condition. The 3D reconstruction of the room in VR fits the shape of the room in the real environment, occluding the real world that needs to be verified in AR.

In the experiment, the tracking capabilities of the Microsoft HoloLens were both applied to the VR and AR conditions, thus, registration error rates should be exactly the same. I also asked participants to stay in one position during navigation to prevent registration errors brought about by moving around, resulting to an unstable tracking. Even when this tracking stability is consistent in both conditions for this experiment, the OST-HMD might not emulate the same tracking conditions of a VR-HMD.

It can also be argued that by solely using Microsoft HoloLens, I am limiting the capabilities of what current VR systems have. Many VR HMDs are capable of wider FoV and rendering quality than the HoloLens. The VR condition in this experiment, then, does not have enough technical specifications to match than that of current VR HMDs. The ideal scenario would be to bring the VR condition in a nearly-real state, minimizing the differences between AR and VR to the point that these conditions are only reduced to analog (real) and digital (CG) representations of the environment.

Notwithstanding other off-the-shelf OST-HMDs (e.g. Meta 2, Daqri), I chose the Microsoft HoloLens for this experiment primarily because it was readily available, and Unity supports building applications on the device. Using the device satisfies the requirements of the experiment system. Even with a small FoV, the Microsoft HoloLens can show the virtual objects clearly and place them in a desired position without too much jitter. It also presented the virtual rendering of the room with enough occlusion for the participants to not see the actual room in the real world.

5.1.3 Other Considerations

Beyond training environment conditions and display device used, the experiment can be extended into other directions. One of which is the skill that is needed to be trained. Memorization and accumulation of spatial knowledge is a fundamental skill. Limiting this study to only one skill provided me with more control in discussing the findings, however, actual training simulators hone many skills at once. I suggest that these kinds of experiments should escalate to higher-order skills like assembly, ordering, and analysis, among others.

This study also did not involve any steps that would involve muscle memory, like walking, or handling of tools. In a practical scenario, most control rooms have workers sit in only one position. Participants can be seated in a swivel chair, with the control room elements within reach. This restricts free body movement to simple swivel rotation of the chair. On another, in cases of larger rooms and unavoidable movement around the space, top-view capture of the room can reveal 2D movement patterns of the user. Future experiments can compare user's sensorimotor performance and training transfer in a task in AR and VR.

5.2 Future Work: Overlay Field of View

The main limitation of the experiment is the fact that we used a Microsoft HoloLens to simulate an immersive VR experience. The small FoV of the HoloLens is very different from the typical immersive VR HMDs that have a large FoV. In a way, this does not fare well with actual VR HMD systems, with a higher resolution and a wider space to view overlay information. As such, the experiment did not illustrate a perfect simulation of an immersive VR experience. While the

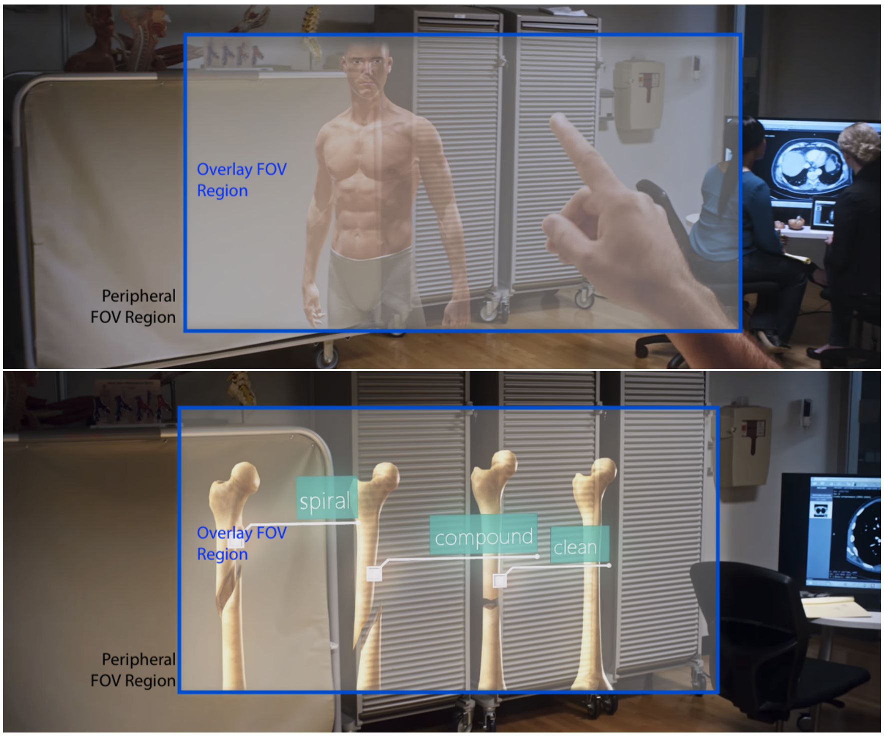


Figure 5.1: Image captures from the Microsoft HoloLens promotional video* annotated with FoV regions

experiment prevented the display of large annotations by just using small indicators fit for the limited FoV region, truncation of large virtual objects can present a perceptual problem. For example, in a visualization application presented by Microsoft*, virtual objects get cut off by the small overlay FoV (Fig. 5.1).

Along with the observation that some participants took a couple of steps back to better observe the scene, the question of how the FoV of the display affects memory transfer should be addressed. For example, Kishishita et al. explored the effects of wide OST-HMD FoV in a multitasking scenario of discovering virtual targets while playing a crossword puzzle [18]. They found that target discovery rates drop as the FoV approaches 100 degrees, and that their two view management setups exhibit a converging discovery rate as the FoV increases from 100-130 degrees.

In the future, I want to investigate how the FoV of the HMD affects the recall

*https://www.youtube.com/watch?v=SKpKlh1-en0 , last accessed 25 July 2018 rate and if the results of our study can be replicated with wide FoV OST-HMDs. I also plan to investigate how other factors, like environment setup affect the memory, and if it could be used in a similar way to an MoL.

Acknowledgements

First of all, I would like to thank the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the screening committee at NAIST for granting me a scholarship under the International Priority Graduate Programs.

I would like to thank Dr. Marc Ericson Santos, Prof. Ma. Mercedes Rodrigo, and Prof. Hirokazu Kato for believing in me five years ago, and selecting me as an intern for one month in the same laboratory that I am in today, the Interactive Media Design Laboratory (IMD-Lab). That experience has led me to choose pursuing my master's degree here at NAIST.

I would also like to thank Assoc. Prof. Christian Sandor and Prof. Kiyoshi Kiyokawa for giving me guidance to improve this manuscript, and to Asst. Prof. Takafumi Taketomi with helping me with the Japanese translation.

I would like to thank the members of the IMD-Lab for sharing valuable input in this study, as I showed my progress and rehearsed my presentation for IEEE VR before leaving. Thank you most especially to Asst. Prof. Alexander Plopski and Jayzon Ty for accompanying me throughout the writing process in order to get the conference paper in its final state.

Many thanks to everyone from the International Student Affairs Section, Academic Affairs Section, and Graduate School of Information Science Office, for the assistance whenever I have problems, and for the immense amount of patience.

I also want to thank the Filipino community at NAIST for being my home away from home these past two years. Finally and most especially, I dedicate this work to my family, for being my support ever since. May we never forget.

"Con el recuerdo del pasado entro en el porvenir."
(Sinasalubong ko ang hinaharap pasan ang alaala ng nakaraan.)
(I enter the future carrying a memory of the past.)
- *El Consejo de los Dioses* (The Council of the Gods), Jose Rizal, 1880

References

[1] Ronald T. Azuma. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4):355–385, 1997.

[2] Timothy Baldwin and J. Kevin Ford. Transfer of Training: A Review and Directions for Future Research. 41:63–105, 1988.

[3] James Bliss, Philip Tidwell, and Michael Guest. The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence: Teleoperators and Virtual Environments*, 6(1):73–86, 1997.

[4] Lisa A. Burke and Holly M. Hutchins. Training Transfer: An Integrative Literature Review. *Human Resource Development Review*, 6(3):263–296, 2007.

[5] Sebastian Büttner, Henrik Mucha, Markus Funk, Thomas Kosch, Mario Aehnelt, Sebastian Robert, and Carsten Röcker. The Design Space of Augmented and Virtual Reality Applications for Assistive Environments in Manufacturing: A Visual Approach. In *Proceedings of the International Conference on Pervasive Technologies Related to Assistive Environments*, pages 433–440, Island of Rhodes, Greece, June 2017.

[6] Andy Cockburn, Carl Gutwin, and Saul Greenberg. A Predicitve Model of Menu Performance. In *Proceedings of the SIGCHI Conference on Human Factors in Computer Systems*, pages 627–636, San Jose, CA, USA, April 2007.

[7] Andy Cockburn, Per Ola Kristensson, Jason Alexander, and Shumin Zhai. Hard Lessons: Effort-inducing Interfaces Benefit Spatial Learning. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1571–1580, San Jose, CA, USA, April 2007.

[8] Nelson Cowan. What are the Differences Between Long-term, Short-term, and Working Memory? *Progress in Brain Research*, 169:323–338, 2008.

[9] Michele Di Donato, Michele Fiorentino, Antonio E. Uva, Michele Gattullo, and Giuseppe Monno. Text Legibility for Projected Augmented Reality on Industrial Workbenches. *Computers in Industry*, 70:70 – 78, 2015.

[10] Andreas Dünser, Karin Steinbügl, Hannes Kaufmann, and Judith Glück. Virtual and Augmented Reality as Spatial Ability Training Tools. In *Proceedings of the ACM SIGCHI New Zealand Chapter's International Conference on Computer-Human Interaction: Design Centered HCI*, pages 125–132, Sydney, Australia, July 2006.

[11] Brian Ehret. Learning Where to Look: Location Learning in Graphical User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 211–218, Minneapolis, MN, USA, April 2002.

[12] Michael Fine and Brandon Minnery. Visual Salience Affects Performance in a Working Memory Task. *Journal of Neuroscience*, 29(25):8016–8021, 2009.

[13] Yuichiro Fujimoto, Goshiro Yamamoto, Hirokazu Kato, and Jun Miyazaki. Relation Between Location of Information Displayed by Augmented Reality and User's Memorization. In *Proceedings of the Augmented Human International Conference*, pages 7:1–7:8, Megeve, France, March 2012.

[14] Yuichiro Fujimoto, Goshiro Yamamoto, Takafumi Taketomi, Jun Miyazaki, and Hirokazu Kato. Relationship Between Features of Augmented Reality and User Memorization. In *Proceedings of the International Symposium on Mixed and Augmented Reality*, pages 279–280, Atlanta, GA, USA, November 2012.

[15] Hind Gacem, Gilles Bailly, James Eagan, and Eric Lecolinet. Impact of Motorized Projection Guidance on Spatial Memory. In *Proceedings of the ACM Symposium on Spatial User Interaction*, pages 51–59, Tokyo, Japan, October 2016.

[16] Lynn Hasher and Rose Zacks. Automatic and Effortful Processes in Memory. *Journal of Experimental Psychology: General*, 108(3):356–388, 1979.

[17] Bonifaz Kaufmann and David Ahlström. Studying Spatial Memory and Map Navigation Performance on Projector Phones with Peephole Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3173–3176, Paris, France, April 2013.

[18] N. Kishishita, K. Kiyokawa, J. Orlosky, T. Mashita, H. Takemura, and E. Kruijff. Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks. In *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality*, pages 177–186, Munich, Germany, September 2014.

[19] Eric Krokos, Catherine Plaisant, and Amitabh Varshney. Virtual Memory Palaces: Immersion Aids Recall. *Virtual Reality*, pages 1–15, May 2018.

[20] E. Kruijff, J. E. Swan, and S. Feiner. Perceptual Issues in Augmented Reality Revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, Seoul, Korea, October 2010.

[21] Eric Legge, Christopher Madan, Enoch Ng, and Jeremy Caplan. Building a Memory Palace in Minutes: Equivalent Memory Performance using Virtual versus Conventional Environments with the Method of Loci. *Acta Psychologica*, 141(3):380–390, 2012.

[22] Svenja Leifert. The Influence of Grids on Spatial and Content Memory. In *Extended Abstracts of the SIGCHI Conference on Human Factors in Computing Systems*, pages 941–946, Vancouver, BC, Canada, May 2011.

[23] Paul Milgram and Fumio Kishino. A Taxonomy of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems*, 77(12):1321–1329, August 1994.

[24] Ulrich Neumann and Anthony Majoros. Cognitive, Performance, and Systems Issues for Augmented Reality Applications in Manufacturing and Maintenance. In *Proceedings of the International Symposium on Virtual Reality*, pages 4–11, Atlanta, GA, USA, March 1998.

[25] Ehren L. Newman, Jeremy B. Caplan, Matthew P. Kirschen, Igor O. Korolev, Robert Sekuler, and Michael J. Kahana. Learning Your Way Around Town: How Virtual Taxicab Drivers Learn to Use Both Layout and Landmark Information. *Cognition*, 104(2):231–253, 2007.

[26] Amir Raz, Mark G. Packard, Gerianne M. Alexander, Jason T. Buhle, Hongtu Zhu, Shan Yu, and Bradley Peterson. A Slice of π : An Exploratory Neuroimaging Study of Digit Encoding and Retrieval in a Superior Memorist. *Neurocase*, 15(5):361–372, October 2009.

[27] Rupert Reif and Willibald Gunthner. Pick-by-Vision: An Augmented Reality Supported Picking System. In *Proceedings of the International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*, pages 57–64, Plzen-Bory, Czech Republic, February 2009.

[28] Oscar Rosello, Marc Exposito, and Pattie Maes. NeverMind: Using Augmented Reality for Memorization. In *Proceedings of the Annual Symposium on User Interface Software and Technology*, pages 215–216, Tokyo, Japan, October 2016.

[29] Valerio Santangelo and Emiliano Macaluso. Visual Salience Improves Spatial Working Memory via Enhanced Parieto-Temporal Functional Connectivity. *Journal of Neuroscience*, 33(9):4110–4117, 2013.

[30] Joey Scarr, Andy Cockburn, and Carl Gutwin. Supporting and Exploiting Spatial Memory in User Interfaces. *Foundations and Trends in Human–Computer Interaction*, 6(1):1–84, 2013.

[31] J. F. Soechting and M. Flanders. Sensorimotor Representations for Pointing to Targets in Three-Dimensional Space. *Journal of Neurophysiology*, 62(2):582–594, 1989.

[32] Md. Sami Uddin, Carl Gutwin, and Andy Cockburn. The Effects of Artificial Landmarks on Learning and Performance in Spatial-Memory Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 3843–3855, Denver, CO, USA, May 2017.

[33] Frances A. Yates. *The Art of Memory*. Random House UK, 1966.

Publication List

[1] Nicko R. Caluya, Alexander Plopski, Jayzon F. Ty, Christian Sandor, Takafumi Taketomi, and Hirokazu Kato, "Transferability of Spatial Maps: Augmented Versus Virtual Reality Training." In *Proceedings of the IEEE Conference on Virtual Reality & 3D User Interfaces*, pages 1–7, Reutlingen, Germany, March 2018.