Doctoral Dissertation

Perceptual Effects of Augmented Reality Experiences Using Head-Mounted Displays on Task Performance

Nicko Reginio Caluya

Program of Information Science and Engineering
Graduate School of Science and Technology
Nara Institute of Science and Technology

Supervisor: Professor Hirokazu Kato
Interactive Media Design Lab. (Division of Information Science)

Submitted on March 17, 2022

A Doctoral Dissertation
submitted to Graduate School of Science and Technology,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
Doctor of Engineering

Nicko Reginio Caluya

Thesis Committee:

Supervisor	Hirokazu Kato (Professor, Division of Information Science)
	Kiyoshi Kiyokawa (Professor, Division of Information Science)
	Masayuki Kanbara (Associate Professor, Division of Information Science)
	Yuichiro Fujimoto (Assistant Professor, Division of Information Science)
	Alexander Plopski (University of Otago, New Zealand)

Perceptual Effects of Augmented Reality Experiences Using Head-Mounted Displays on Task Performance[*]

Nicko Reginio Caluya

Abstract

Continuous improvement of task performance through assisting cognitive processes may find benefit in virtual and augmented reality (VR, AR) environments, deployed in head-mounted displays (HMDs). Through HMDs, users can view virtual objects much closer to the eyes. These objects can be task stimuli superimposed onto the real world or onto its approximate replication through mixed reality (MR) simulations. As tasks rely on details of the real world, AR experiences shown on HMDs have the potential to impact the performance of users. However, current HMD designs show only a limited amount of information that can be rendered within its view, which poses perceptual challenges.

The main goal of this dissertation is to investigate different visual factors of simulations deployed in HMDs that may affect the performance of tasks. I conducted two experiments: one involving spatial memory and the other involving speed perception. The results of the first experiment showed that the overlay field of view (OFOV) size of HMDs used in AR did not affect spatial memorization, but wide OFOV reduced head rotation. The results of the second experiment showed that the information density of augmented vection patterns within the optical see-through HMD view influenced driving speed perception, however, other visual properties still need to be considered and modified to fully realize a speed perception-based AR. The findings of these experiments could have implications on the design and HMD choices of augmented training and related experiences.

---

[*] Doctoral Dissertation, Graduate School of Information Science, Nara Institute of Science and Technology, March 17, 2022.

Keywords:

augmented reality, spatial memory, head-mounted displays, field of view, speed perception, information density

Contents

List of Figures

List of Tables

1. Introduction

When users experience augmented reality (AR), they view digital information (e.g., virtual objects in 3D) that is superimposed against the real world. These users can view such information through a variety of systems and devices that may be handheld, head-mounted, or projected. In recent years, the availability of consumer products that support AR-based experiences have risen, and with variety comes ergonomic and technical benefits that make AR a useful technology for many tasks.

1.1 Augmented Reality Using Head-Mounted Displays

The devices used to display AR vary from one context to another. Kruijff et al. [41] enumerated different perceptual issues in AR, in which some were associated with three major classifications of AR devices: head-worn display (which I will refer to throughout this dissertation as head-mounted displays or HMDs), mobile devices, and projector-camera systems.

Spatial AR leverages the use of projectors and projector-camera systems to adjust lighting and rendering conditions. Handheld AR takes advantage of portable and mobile devices like tablets and smartphone to serve as the display interface at the expense of limited manipulability of the interface, as the user holds the device while immersed in the experience.

In this dissertation, I focus on experiences deployed on HMDs, which have several advantages. With the availability of new off-the-shelf HMDs, users can now move around the area, and untethered HMDs also expand the area in which AR can be experienced. Most importantly, the viewing area is directly and closely within the human field of view, which makes visual perception of rendered virtual objects easier and more convenient with HMDs than with mobile or projection-based devices.

1.2 Human Visual Information Processing

Humans perceive visual stimuli and integrate various cognitive processes. In the model of human information processing stages (Fig. 1.1), prior knowledge

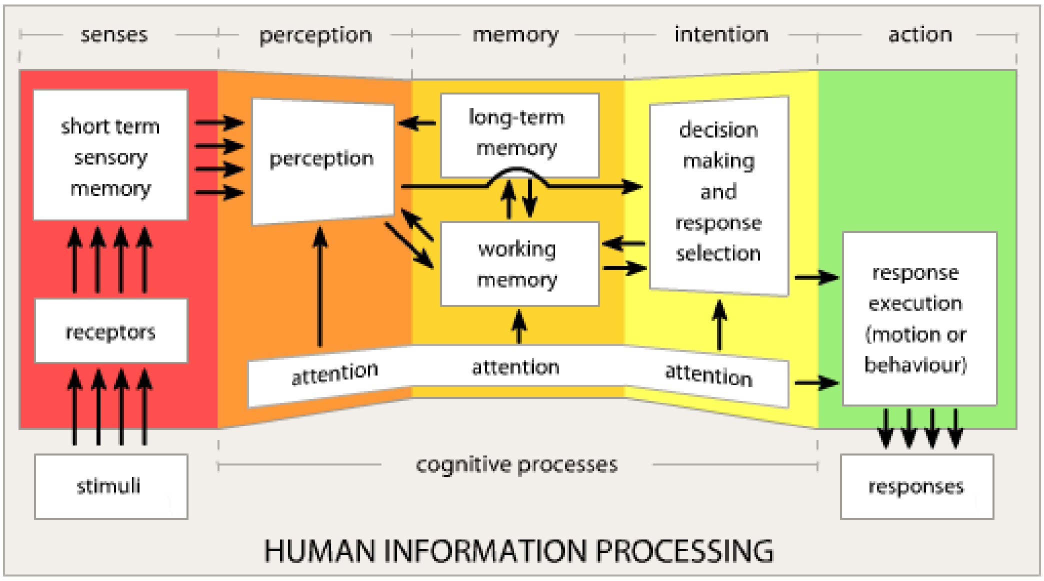


Figure 1.1: Illustration adapted from Wicken's model of human information processing [72].

stored from memory integrates with sensations to perceive stimuli. This dynamic interaction among sensation, memory, perception, and attention all contribute to the decision-making processes and response selection producing feedback.

Research on human-computer interaction (HCI) technologies that complement fundamental cognitive processes contribute to further understanding of ways humans process information and perform tasks. As AR experiences overlay information that we do not normally perceive in the real world, I see the advantage of having visual augmentations appear alongside real world stimuli. As OST-HMDs are wearable, its close proximity to the eyes provides an ergonomic benefit by overlaying the new layer of information immediately.

The experiments that I present in this dissertation shed light on two cognitive processes: spatial memorization and speed perception. I will be defining these two processes in more detail in Chapters 2 and 3, respectively.

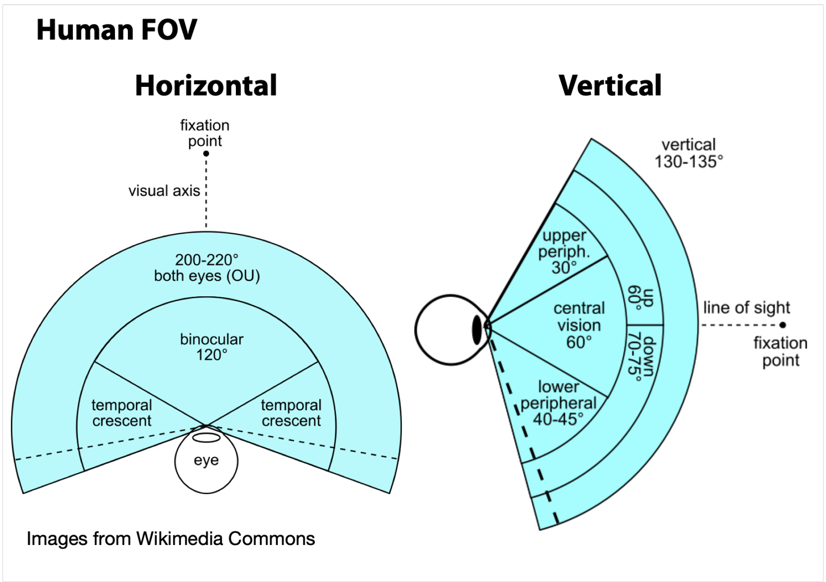


Figure 1.2: The cross-sections of the human visual field: horizontal and vertical field of view

1.3 Limited Amount of Information in HMDs

1.3.1 Field of View in Humans and OST-HMDs

Viewing such information should suit our human eyes, which have a binocular field of view (FOV) at 114° horizontally [31] (see Fig. 1.2). Aside from the conventional horizontal and vertical FOV visualizations, many studies have also opted to map the FOV in polar plots, or represent the FOVs diagonally (see Fig. 1.3). For the purposes of this dissertation, the FOV sizes are indicated as horizontal, unless otherwise stated.

Off-the-shelf HMDs offer different FOVs. In particular, optical see-through HMDs (OST-HMDs) have significantly smaller FOV than immersive (non-see-through) HMDs, which means less space to augment information or virtual objects. To clarify, the FOV referred to here is the overlay FOV (OFOV), different from the FOV of the world, which is usually much larger (see Fig. 1.5).

For example, Microsoft HoloLens has 34° and 52° diagonal OFOV for the first

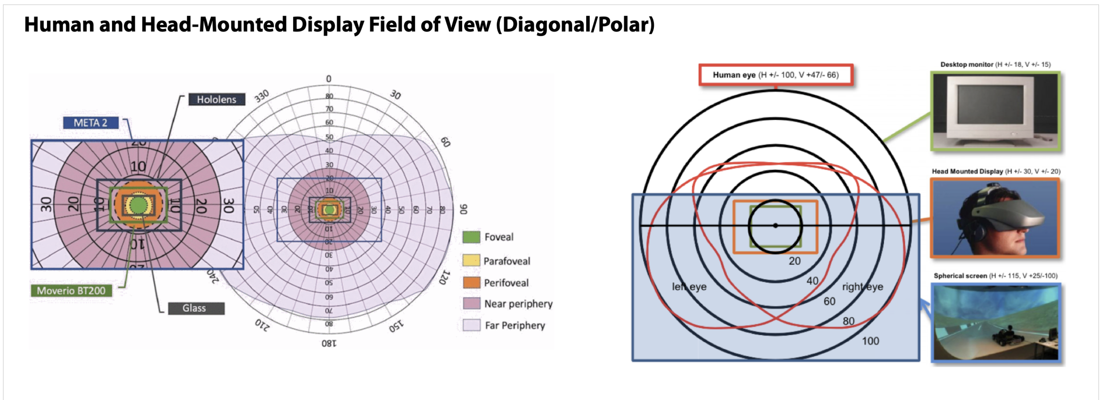
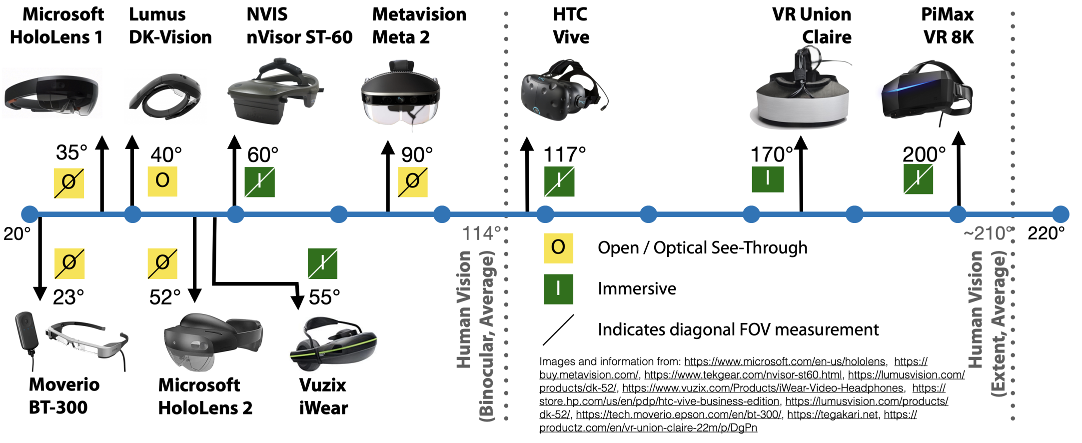


Figure 1.3: Polar plot representations of the FOV by [55] and [70]



Images and information from: https://www.microsoft.com/en-us/hololens, https://buy.metavision.com/, https://www.tekgear.com/nvisor-st60.html, https://lumusvision.com/products/dk-52/, https://www.vuzix.com/Products/iWear-Video-Headphones, https://store.hp.com/us/en/pdp/htc-vive-business-edition, https://lumusvision.com/products/dk-52/, https://tech.moverio.epson.com/en/bt-300/, https://tegakari.net, https://productz.com/en/vr-union-claire-22m/p/DgPn

Figure 1.4: A sample range of off-the-shelf HMDs with their respective FOV sizes.

and second generations, respectively[1], while Meta 2 has a 90° horizontal OFOV[2]. On the other hand, immersive HMDs have a much larger FOV, e.g., 110° of the HTC Vive [3] or 200° of the Pimax Vision 8K X[4].

---

[1] https://web.archive.org/web/20200922182626/https://www.wired.com/story/microsoft-hololens-2-headset, accessed 28 May 2021

[2] http://web.archive.org/web/20181120212212/http://buy.metavision.com:80/products/meta2, accessed 19 August 2021

[3] https://web.archive.org/web/20210516235847/https://www.vive.com/eu/product/vive/, accessed 19 August 2021

[4] http://web.archive.org/web/20210812021035/https://pimax.com/product/vision-8k-x/, accessed 19 August 2021

1.3.2 Information Density in OST-HMDs

The FOV of OST-HMDs also limits the density of information that can be displayed, due to the hardware design or trade-off towards better pixel quality of the rendered virtual image. While this issue is usually attributed to screens and interfaces outside of the context of AR, there have already been studies that investigate information density of OST-HMDs as a factor. Some looked at the relationship between information density in visual search performance [70], or information management techniques for virtual objects that are outside of the effective OFOV [36].

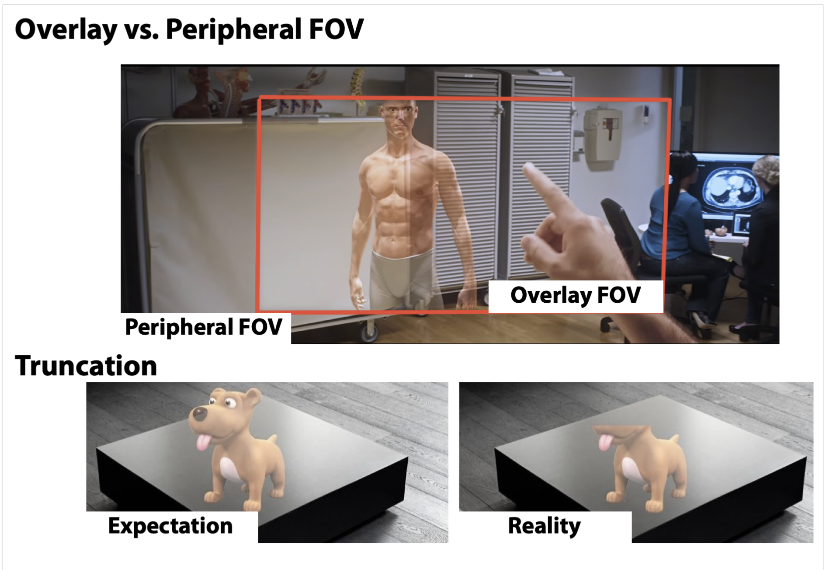


Figure 1.5: The truncation of virtual renderings due to small overlay (OFOV) against the peripheral FOV (PFOV)

1.4 Contributions and Dissertation Overview

I explore two tasks that exploit fundamental phases in visual information processing: memorizing a spatial map of the room and perceiving one's own driving speed.

The addition of HMDs in the experience presents a unique visual experience. For AR HMDs, particularly in OST-HMDs, an additional layer lies between the eye and the real world, and such a layer has superimposed information. In the context of training, the content to be learned in this case is presented as virtual objects rendered in a holographic display. For a closed-view HMD in VR, the actual view of the real world is absent. The device displays both the immersive environment and the training content as virtual renders.

This dissertation presents two empirical experiments. First, I designed a simulation of mixed reality (MR) to change the sizes of the OFOV of OST-HMDs, and investigated how the OFOV affects spatial memory. Second, I designed an animated pattern with varying information densities displayed on an OST-HMD, conducted an experiment to investigate their effects on driving speed, and drew insights about the perceptual process of visual speed.

I would like to contribute to the growing research of human factors that involve the use of HMDs, specifically in OST-HMDs. My primary motivation for this research is the creation of HMD-based augmented training solutions that are efficient and effective in delivering training content to trainees. Furthermore, I would like such solutions to facilitate a faster transfer of spatial skills, such that the trainees will need minimal dependence on HMDs while maximizing the benefits of training on one, so that they can be able to perform the task in the real world correctly and efficiently. Additionally, I would like to explore the relationship between the features or hardware limitations of current (OST-)HMDs, the task performance of users, and the cognitive processes that are involved when these scenarios are designed.

As this dissertation comprises two major experiments, I organized the properties in Table 1.1 to show the differences between them.

Table 1.1: Differences between the two major experiments in this dissertation

Properties	**Chapter 2**	**Chapter 3**
Cognitive Process	Spatial Memory	Speed Perception
Ind. Variable (IV)	Overlay FOV	Stimulus Density
Dependent Variable (DV)	Test Scores (Recall, Transfer)	Actual Driving Speed
AR Content /Visual Stimuli	Pair of targets in fixed locations	Vection pattern with outward radial motion
Relationship with the OFOV	Truncated if outside OFOV	Visible within OFOV bounds /periphery
Visibility on Central View	No need for out-of-OFOV visualization	No visible stimuli at center
Virtual Environment	Semi-circular layout of virtual panel boards	Simulated driving scenario
Real World Environment	Semi-circular layout of actual panel boards	Driving seat and semi-circular monitor setup
Device	VR HMD (OST-HMD Simulation)	OST-HMD (Simulated Driving Scene)

2. Effects of Field of View of Head-Mounted Displays on Spatial Memory

One of the main targets of criticism of HMDs is the FOV size, whether in VR or AR. This limitation is prominent with OST-HMDs, as those with narrow overlay FOV (OFOV) sizes only provide a small window to view virtual objects. I want to investigate if restricting this OFOV negatively affects a user‘s ability to memorize spatial locations in a simulation of a work environment, and consequently, long-term memory transfer to an equivalent scenario in the real world two days later. I conducted a within-subjects experiment with 18 participants, performing in three phases with an OST-HMD, simulated on an immersive HMD. For each phase, they viewed the training scenario with a different FOV size of the augmentable area (30°, 70°, 110° diagonal).

Results from recall tests showed that smaller FOV size did not significantly affect user’ s performance on both short-term and transfer tests, but HMD data revealed that users rotated their heads less with a 110° OFOV. Furthermore, proximity of objects to memorize had an interaction effect with smaller OFOV sizes. The findings of this study point to the consideration of various aspects of training spatial skills in AR and properties of the HMDs used for such training.

2.1 Introduction

Current technologies like virtual and augmented reality (VR, AR) have been used to facilitate training sessions by simulating the work environment [76, 24], in case of inaccessibility of the workspace or lack of training space. Skills like spatial awareness and understanding rely on the quality of spatial information, which makes AR or VR viable training platforms. For example, astronauts need to be aware of the changing spatial orientation as they perform tasks under the influence of lower gravity [52]. Some tasks require memorizing the layout of the space, then transferring the retained information in another form and location [2, 53], potentially reallocating time from memorization to other critical operations. The visual designs of AR/VR environments and appropriate placement of task information in the scene should then be considered to support positive training transfer (i.e., continuous and efficient performance [7]) to the actual task.

2.1.1 Purpose and Overview

Since different HMDs provide different FOV sizes, I would like to investigate whether the OFOV size affects the memorization of spatial relationships. Various studies showed that limiting a user's OFOV will negatively affect their ability to memorize the layout of the environment [2, 3] and time to navigate the environment [68]. Intuitively, one would expect a small OFOV to, similarly, negatively affect the user as it truncates the view of the virtual objects. The limited OFOV could also lead to partial loss of information that lies outside it, further hindering the learning process. However, this effect could be counteracted by the increased focus users need to exert to memorize the locations when they are not visible all the time under a smaller OFOV. Furthermore, users focus primarily on the object at the center of the visual field [64], limiting the benefits a larger OFOV could have on the memorization task. I specifically chose memorizing two virtual objects because their distance would affect the way they are viewed with an HMD with a restricted FOV. I see the benefit of exploring this task on operations like gaze-free control, faster search, and cross-checking of spatial information. Furthermore, some operations in the real world involve both immensity and complexity like circuitry and cable work, where memorizing the connection between two points is necessary.

This chapter discusses the results of a within-subjects experiment ($N = 18$) that investigates the effects of the size of the OFOV on a user's ability to memorize locations of interest. As I did not have access to a large FOV OST-HMD, I utilized a simulation of an OST-HMD in VR that allows us to explore FOVs not currently possible on most commercial OST-HMDs. Such simulations of AR have been shown to provide results similar to those conducted on the actual device [43] and have been used to study navigation techniques for OST-HMDs [61]. I simulate three OFOVs (30°, 70°, 110°, measured diagonally, as seen later in Fig. 2.1). For each OFOV participants had to memorize the locations of 15 target pairs over the course of four memorization tasks. Each memorization task was followed by a recall task to monitor the participants' progress. Finally, I evaluated retention on a transfer task two days later in a real environment resembling the virtual training scenario (Fig. 2.3). The results show no significant difference among the OFOV sizes in terms of correct responses during the recall tasks as well as the transfer task. However, the recorded HMD rotation data revealed that participants had the least amount of overall rotation in the 110° condition, which lessens potential physical strain and reduces overall memorization time.

2.1.2 Research Contributions

The contributions of this chapter are as follows:

- to conduct an empirical study and gather subjective feedback about the relationship between OFOV in OST-HMDs and spatial memory,
- to provide insights about the relevance the OFOV of an HMD has in influencing the way a user perceives and carries out a task, particularly highlighting the issue of incomplete perception of the virtual objects even when the real world is fully shown in an OST-HMD setup.
- to highlight observations made on spatial memorization that could have implications on the design of HMD-based simulations as training environments (e.g., trainees exert less head rotation on large OFOVs, the proximity of virtual objects as training stimuli is important), and
- to present results and findings relevant to the discussion of whether or not designers should be wary about the OFOV to support the training process.

2.2 Related Work

Applications in AR and VR have demonstrated their potential of assisting many spatial tasks and improving the spatial skills of users, such as spatial memorization. In many studies, however, the FOV of the AR and VR devices have been commonly cited as a factor affecting the development of such applications and the users' performance.

This section reviews prior explorations of AR and VR for spatial memorization, followed up with a discussion of effects a limited FOV could have on users and ways to mitigate it.

2.2.1 Classifications of Memory

The time of accessing stored spatial memory (and memory in general) is important. The common classification of memory is temporal: as short-term and long-term memory. During successive and immediate iterations of exposure to memorized stimuli, working memory should be able to build short-term memory. Then, as a significant amount of time elapses, the stimuli used to memorize are no longer present, or the environment changes, memory may undergo deterioration. Memory that still remains at this time can be referred to as long-term memory.

This distinction between the short-term and long-term memory is also one of the reasons behind the design of the experiment. Not only do I want to investigate immediate effects, but I also want to extend the duration of the study so that I can also see what happens long after the training has finished.

I would like to highlight that the long-term memory that I am exploring in this study is unique. The environment used during the training and the one used to evaluate spatial memory is different in terms of realism. I would like to situate the participant of the experiment in the real world, where the elements of the room resemble that of the training environment in the virtual world, thus demonstrating a training transfer scenario.

Aside from duration, the types of memory can also be organized according to heuristics. In this taxonomy, memory can be classified as declarative (explicit) and non-declarative (implicit) [67]. Within declarative memory are two common types of memory: episodic and semantic. Episodic memories involve events that have happened to a person, replayed as snapshots or short episodes. On the other hand, semantic memory involves recalling meaning and general knowledge.

For the purpose of this experiment, I define spatial memory in this experiment as the ability to recall spatial information (e.g., location) of objects. I wanted to make this distinct from episodic and semantic memories, such that the memorization task should not depend on the sequence of events as in an episode, or on the meaning of the stimuli or object (e.g., no words or icons).

2.2.2 Spatial Memorization in AR and VR

Critical operations such as fast access and manipulation of information (e.g., flight control [14]) and navigating hazardous spaces (e.g., astronauts missions in space [52]) require good spatial memory. However, such critical operations may need to be rehearsed in a different environment, and hazardous spaces should be simulated with safety precautions. This is where training environments through simulations in AR and VR can be helpful.

Benefits of AR and VR to learning have been studied extensively, showing that both can support memorization [9, 22, 21, 62, 40]. It is important to note that the amount of time passed before short-term and long-term recall tests varies depending on the study [22, 32, 15, 5, 4]. Generally, short-term memory recall is considered to happen immediately after a memorization task (within 30 seconds) while long-term recall refers to any event where the memory is probed at a later time.

The efficiency of the learning process can be affected by different parameters, such as the number of dimensions of the environment [11], the number of repetitions and items to remember [4], or the presence of landmarks [71, 23], which was found to be more beneficial for learning the location of elements than a grid layout.

Although it is more difficult to memorize items distributed in a 3D space over 2D [11], the 3D nature of the experience has been shown to be beneficial for the loci memorization technique [62, 40]. This technique associates locations the user is familiar with to pieces of information. When trying to memorize the information, the user creates a mental walkpath between the different locations, recollecting the information as they continue on the path. This integration of spatial information can assist the retention process.

Others explored how directly augmenting the environment by overlaying the content associated with the location onto it, could be beneficial for the learning process. Gacem et al. [22] compared the effects of different confirmation and highlighting techniques on learning the spatial location of elements around the user, but did not investigate the FOV. The experiment from my master's thesis [9] compared learning of spatial locations in AR and VR, and found that VR resulted in better short-term recall, while AR led to better long-term memorization. Additionally, although participants were asked to stand at a pre-defined location, they tended to step backwards in both environments to increase their view of the world. One potential explanation for this behavior is the limited OFOV of only 30° in both conditions. Contrary to these works I want to investigate the effects of different OFOV sizes, while the user has an unconstrained view of the scene.

2.2.3 Effects of FOV Restriction

Constraining the user's FOV has been shown to result in reduced performances in a variety of tasks related to spatial awareness, such as search [18, 70], inspection [59], navigation [68, 8, 30], driving and flying [13, 14], as well as retention of the explored environment [34, 49, 2].

However, as noted, by Lin et al. [49], this is not universally the case. For example, in their study increasing the FOV beyond 100° did not result in improved memorization of a space while driving. Supporting this observation, Knapp and Loomis [39] found that a smaller FOV did not significantly affect depth estimation, while Hassan et al. [30] showed that small FOV angles could be sufficient to navigate a scene with high contrast. Ragan et al. [58] found that a larger FOV significantly improved training results of a search task in VR, but did not result in a significant difference during assessment. Similarly, Kishishita et al. [36] found that when two tasks (search and puzzle solving) compete for attention, a larger FOV did not lead to an improvement in search performance. These observations show that although generally a larger FOV results in improved performance, it ultimately depends on the task at hand.

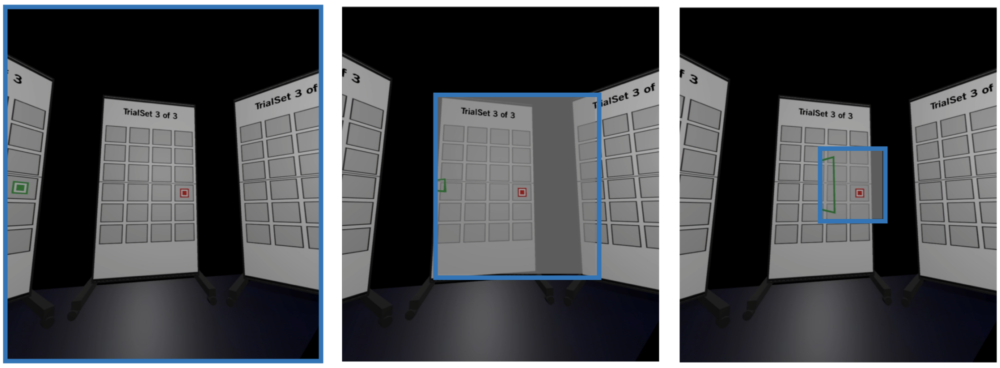
Several techniques have been explored to reduce the aforementioned negative effects of the FOV restriction by the HMD housing. Some techniques mitigate it by incorporating LEDs and low-resolution displays to provide information about the occluded areas [56, 28, 74]. Others approached this as a view-management problem and indicate the existence of occluded objects via in-view or in-situ labeling [36], or screen edge indicators [56, 26]. Both approaches have been shown to effectively assist users in maintaining awareness of objects occluded by the housing of the HMD, however indications with CG bear the benefit of being easily deployable on existing hardware, without need for customization.

2.2.4 Summary of Related Work

The review of related work showed that although many papers explored the effects of the FOV on user performance, these usually focused on reduction of the overall FOV, not only the augmentable portion of the FOV. If the user's view of the environment remains the same, their information of the scene layout is not affected by the OFOV. As such, users could memorize the information presented on the OFOV relative to the environment, reducing the benefits a larger OFOV. While different studies have considered the benefits of different parameters in AR and VR on memorizing locations, they did not investigate what effects a limitation of the OFOV would have on the results. As such, the goal of this experiment is to determine if the intuitive assumption that a larger OFOV size leads to better memorization results truly holds. The tabular summary of related work can be seen below (Table 2.1), showing the different viewing displays or devices used, the levels of FOV used in the experiment, the number of participants, and the treatment used in each study design.

Table 2.1: Positioning of this Spatial Memory Experiment Among Related Work

Ref	**Viewing Display/Device**	**IVs (FoV Sizes); H - Horizontal, V - Vertical**	**Subjects (#)**	**Treatment**
[2]	Goggles	9°, 14°, 60°, Full	54	between
[39]	Cardboard-based Goggles	Unrestricted (180° x 120°), Restricted (47°x 43°)	10	within
[36]	Wide-FOV OST-HMD	VxH: 36° x 20.3°, 54° x 30.4°, 81 x 45.6°, and 100° x 45.6°	16	within
[18]	CAVE	8, 16, 32, 64, 128°, SNR=No Restriction	24	between
[68]	Pilot goggles with green filter	FOV: 20°, 40°, 60° and view type: Monocular, Binocular	12	within
[68]	Ski goggles with black cardboard masks	H: 30°, 75°, 112°, 120°, 140°, 160°,180° x V: 18°,48°	12	within
[60]	Allosphere, PhaseSpace Impulse X2	Full (∼220°x120°), (45°x 30°)	30	within
[55]	Panoramic Quarter Sphere Screen	H: 30°, 60°, Full = No restriction	10	within
[55]	Panoramic Quarter Sphere Screen	Central and Peripheral Occlusions: 10°, 20°, 40°, 60°, Full	10	within
[20]	Oculus Rift DK2	Internal-External FOV: 120° – 155°, 68° – 120°, 58 – 110°, 50° – 100°, 43° – 90°, 36° – 80°	30	within
[44]	HoloLens 1	Unrestricted (∼220°x120°), Restricted (30°x 17°)	26	within
[30]	VR HMD with 53°x41° FOV	Diameter: 10°, 20°, 40°, Contrast levels: low(3%), medium (6%), high (11%)	20	within
This	**HTC Vive (MR Simulation)**	**Diagonal: 30°, 70°, 110°**	**18**	**within**



110° : Both targets are visible

70° : Red target is visible, green wedge indicating the green target is close

30° : Red target is visible, green wedge indicating the green target is far

Figure 2.1: The participant's egocentric view (left-eye camera) while performing the memorization task, in three different levels of field of view ($OFOV$) (a-c, blue outline was not visible to the participants). When combined with the right-eye camera, the aspect ratio of the full view doubles from 0.9 to 1.8.

2.3 Experiment

The goal of this experiment is to compare the effects of the OFOV size on memorization of spatial relations in a virtual reality environment and its transfer to the real world.

I focused on the visual perception of users while wearing an immersive HMD when discussing the OFOV. When this OFOV size changes, the available contents in AR that can be seen without moving the viewport also change. I hypothesized that the FOV will have no impact on memorizing a single location consecutively as we tend to focus on that location. This meant that the size of the OFOV would at best impact search performance. As such, I considered a task where the larger OFOV could present a benefit, such as circuitry work. Thus, I designed a spatial memory test viewed on an HMD where the OFOV size can be changed.

I considered the short-term performance of users when exposed to many and successive iterations of memory stimuli. Previous studies [22, 9] showed that participants memorized more locations over multiple iterations of learning the locations. Second, studies mentioned in related work conducted memory tests immediately after and in the same location. I want to further explore the reten-

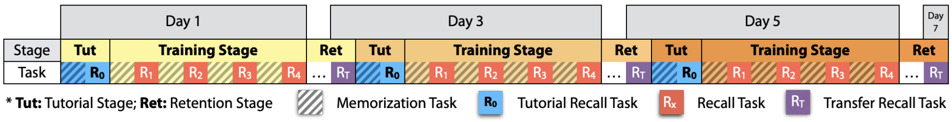


Figure 2.2: Experiment procedure showing the order of events for one participant over the course of seven days. Each color on the Stage row represents a different phase of the experiment. Furthermore, the levels of $OFOV$ are also different for every tutorial and training stage.

tion of memorized information through longer time (i.e., days after) and through relocation (i.e., training transfer to the real world).

The primary instruction for the participants was to memorize the location of 15 target pairs shown during the *memorization task* and recall their location in the following short-term *recall task*. I chose 15 as the number of items based on preliminary tests, presenting a challenging task which helped in avoiding a potential ceiling effect. Thus, I did not expect participants to memorize all items. The memorization and recall tasks were all conducted in the virtual environment. I was interested in the long-term retention and transfer of the learned content, as knowledge gained in VR may not be fully applicable when transferring to a different environment [47]. I asked participants to take part in an additional *transfer task* conducted 2 days later on the physical counterpart of the virtual environment. The transfer task was conducted similarly to the recall tasks and also had a maximum score of 15. Participants performed the memorization and recall tasks with an immersive HMD, where I modified the view with a different OFOV size for each day (Fig. 2.2).

2.3.1 Hypotheses

When I designed the experiment, I formulated the following hypotheses about the impact of OFOV size would have on some aspects of efficient training: the speed of the training, subsequent mastery of the skill, the economy of trainees' motion, and the satisfaction of the trainees when using the training application.

H_1: Smaller OFOV sizes will result in lower scores in recall and transfer tasks.

For H_1, I compare the test scores between OFOV sizes. As experiments on spatial memory in VR involve training transfer to a different environment [53, 57], I made a distinction between the repeated training within the same environment (HMD) and the transfer to the real world 2 days after. Thus, this hypothesis can be divided as follows.

H_1A: **Smaller OFOV sizes result in lower test scores in short-term recall tasks.**

As the trainees can view the training objects more fully on large OFOV sizes, they can view two target pairs simultaneously with ease. Furthermore, I hypothesize that trainees may be more prone to error and mislocate the missing targets when the view is truncated or discontinued when the OFOV sizes are smaller.

H_1B: **Smaller OFOV sizes result in worse retention, leading to lower scores on the long-term transfer task.**

In a similar argument, targets memorized in smaller OFOV sizes would have a weaker mapping of the target pairs because the trainees would only see them partially. Consequently, this weaker mapping would mean key information would be missing (e.g., forgetting at least one of the target pair's locations), which would negatively affect training transfer.

H_2: **Target pairs that are close to each other are better memorized, thus, scores coming from adjacent target pairs are higher than far target pairs.**

Controlling the distance between the two targets contributes to the necessity of having a specific OFOV size. I hypothesize that small OFOVs may not be enough to completely view two *Far* targets conveniently without moving one's head. On the other hand, target pairs like *Adjacent* ones may register immediately making the memorization process faster. Thus, I suggest a secondary variable of *Proximity* between target pairs as an additional factor for investigation.

I assume the possibility of interaction between *OFOV* and *Proximity.* In particular, I hypothesize that the poorest recall performance would happen when participants attempt to memorize *Far* target pairs with the smallest *OFOV* size in the experiment, 30°.

H_3: Smaller OFOV sizes result in longer recall times when answering short-term recall tests and transfer tests.

Connected to the first hypothesis, I assume that there will be more difficulty in recalling target pairs that were presented in smaller OFOV sizes. In this context, however, this difficulty manifests in the longer time it takes for participants to come up with a final answer, regardless of the correctness of the answer.

H_4: With smaller OFOV sizes, participants will perform slower and require more effort.

This experiment also looks into time and speed as a measure of performance. In this hypothesis, I would like to identify two particular segments of the experiment process: the memorization and recall task completion time during the training phases and the amount of head rotation when performing the memorization task.

For the memorization and recall task completion time I assumed that trainees would require more time when viewing smaller OFOV sizes to completely register the target pairs as visual stimuli. I also hypothesize that the speed at which participants rotate their head and shift their view to their left or to their right (i.e., y-axis rotation) would be higher for smaller OFOV sizes.

2.3.2 Experimental Environment

The motivation for this experiment came from previous studies on search and memorization. In particular, I constructed a curved wall layout, similar to the one used by Gao et al. [23]. The main benefit of the curved layout over the planar layout used by Gacem et al. [22] and my previous work [9] is that the relative size of the target is not affected by its location on the surface, i.e., targets further away from the participants do not appear smaller than those right in

(a) Training Environment. (Virtual) **(b) Transfer Environment.** (Real)

Figure 2.3: The training was conducted in a virtual environment (a) with a blank background and floor, while the transfer environment (b) in the real world used a green curtain backdrop and floor.The panel boards were visible in the same way across all three OFOV sizes.

front of them. While some studies showed that utilizing landmarks helps the memorization process [71] I used a simple grid layout to avoid biasing the results. While locomotion paired with a bigger visual overview of the environment may benefit spatial memory [34], I opted not to ask participants to walk around the environment to avoid motor memory from affecting the results. This also allowed the exclusion of potential confounding factors from different OFOV sizes affecting how efficiently participants navigated the environment.

I built the curved wall from five panel boards that were equally spaced in a semi-circle around the user seated on a stationary chair (40 cm height). They appear as 3D models in the virtual world and actual panel boards in the transfer environment (Fig. 2.3). Each panel board (166 cm x 94 cm) contains 6 rows x 4 columns of 16 cm x 16 cm square icons inside it, shown as two 4x3 grids, one at the top and one at the bottom separated by a thick black line that serves as a divider. Each square in the grid only has a black border and a white background.

I kept the environment minimal to allow participants to focus on the memorization task and to avoid memorable objects functioning as landmarks. This allowed the close replication of the virtual environment in the real world, but allowed for some differences requiring users to apply their knowledge in a different environment. In the virtual environment the panels stand on a blue floor in front of a gray background. In its physical counterpart I use a green curtain and floor as the backdrop.

2.3.3 Implementation

Mixed Reality Simulation In the development of the application that I show to the participants, I needed to create a close replication of the real world environment so that I can make a proper simulation. I measured the dimensions and positions of the five panel boards as accurately as possible. Then, I created 3D models using the measured information. Then, I exported these files onto Unity 2018 LTS, where I did the development of the application.

To distinguish between the virtual objects that need to be memorized (V) and the simulation of the real world in the virtual environment (S), I created a stencil shader. This shader renders only the portions of V within the bounds of a viewing frustum simulating an OST-HMD with a specific OFOV level. I add another plane in front of the main camera, acting as a viewing area that simulates an OFOV. On all the levels of $OFOV$, the rendering of S is unaffected, as the main camera renders everything else.

Wedge Visualization While the goal was to keep the two environments similar, participants were assisted in the virtual environment by indicating the target locations relative to their current view. This prevented the task from becoming a search task and helped participants to efficiently explore the virtual environment. I opted to recreate the 3D wedge indicator described by Gruenefeld et al [26] (Fig. 2.4) in my application. This technique does not require physical modifications of the HMD frame and has been shown to be effective in guiding users. The wedge is a triangle that has one of its corners at the target location and whose opposite side sticks into the OFOV. The wider the visible side of the triangle, the further the participant needs to rotate their head to bring the target into their view. The 3D wedge thus encoded not only the direction, but also the distance to the target making it easy for participants to find it. I rendered the wedge in the same color as the target it was pointing to. During training I showed the wedge for both targets, while only the wedge for the target shown to the participants was visible during the short-term recall tasks (Fig. 2.1).

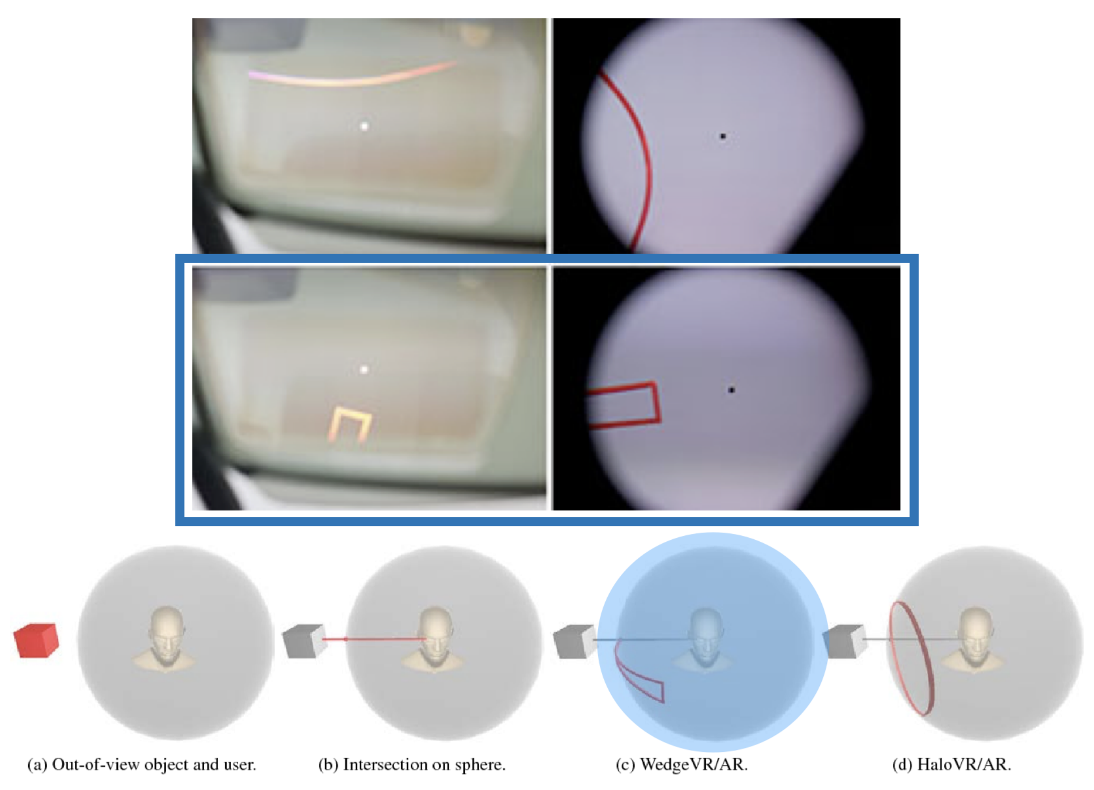


Figure 2.4: The 3D Wedge and Halo visualizations for AR as originally illustrated in [26], with the 3D wedges I highlighted in blue.

2.3.4 Experiment Variables

The independent variable of this experiment is the *OFOV* with three levels (30°, 70°, 110°). The number of times I assessed participants for each level are also defined. The names of the variable of short-term recall tasks during training stage are R_1-R_4 and the transfer task after 2 days is R_L, respectively.

From the review of prior work I could not identify a commonly used set of OFOV sizes to evaluate its effects on performance (see Table 2.1). Usually, the maximum FOV size is determined by the used device, and any additonal levels are spread between an arbitrarily set minimum and this upper threshold [70, 36, 3].

I used the HTC Vive[5], an immersive HMD capable of displaying up to 110° diagonal FOV, which became the upper threshold I could investigate. I selected

---

[5] https://www.vive.com

a diagonal FOV of 30°, which is slightly smaller than the OFOV of the Microsoft HoloLens used in my master's thesis [9], as the smallest FOV. Finally, I added an intermediate value of 70° to observe the OFOV variable in three levels. The different OFOV sizes used for the experiment affect the appearance of virtual objects when they are positioned in the periphery (Fig. 2.1).

As an additional variable to consider, I included the *Proximity* of each target pair. I defined 3 categories of proximity according to the panels: *Adjacent*, *Near*, *Far*. An *Adjacent* pair of targets are located on the same panel or at immediately neighboring panels. The panels of *Near* targets have one panel in between them. Finally, the *Far* target pairs are on the extreme opposites of the setup, having two or three panels in between them. Since the vertical distance is limited (to at most 5 squares apart), I have randomized this distance for the targets. Nevertheless, each target pair in a set has only one matching target pair in another set in terms of the X-Y distance and proximity category.

I counterbalanced variables in the experiment, mainly with the three *OFOV* sizes, using all order permutations. To avoid carry-over effects between different OFOVs I defined three unique target *Set*s (Fig. A.1), consisting of 15 target pairs each. To ensure that all target *Set*s present a similar degree of difficulty I first defined a single *Set* that satisfied the *Proximity* requirements and mirrored it horizontally and vertically to generate the other two *Set*s. For each *Set* I randomly assigned the color of the targets (i.e., red or green). I also made sure that the resulting *Set*s do not overlap. To avoid confusion between learning difficulties of a particular *Set* and the *OFOV* levels I counterbalanced the combinations of $OFOV \times Set$.

The dependent variable that determines the training performance of the participant is the test score, measured as the number of successful trials during the recall and transfer tasks with a maximum score of 15. Aside from the score, I have also included the performance time of the participants in the analysis. I want to observe the duration of the training stage in terms of memorization task completion time, the recall times on all the recall tasks, as well as the overall training time for each OFOV level. As I restricted the movements of the participant to staying put in one location, I also gathered rotation data from the HMD to observe movement differences across all *OFOV* levels.

I evaluate the subjective workload of the system under different OFOVs with the NASA-TLX Questionnaire [27]. I also collected qualitative feedback about the participant's experience.

2.3.5 Experimental Procedure

I recruited 18 participants from our university through an online announcement (5 female and 13 male, 24.78 ± 2.37 years of age). Based on a pre-experiment survey, only 5 participants used an HMD for the first time. While 13 out of the 18 participants reported the use of personal visual aids for daily use (e.g., contact lenses, eyeglasses), they did not experience any difficulty viewing the scenario in the VR HMD. While performing the experiment, these participants had normal or corrected-to-normal vision and reported no other visual impairments when I asked them before proceeding.

The experiment took place over the course of 7 days, as shown in Fig. 2.2. It comprised of 3 phases spanning 3 days each. A new phase started immediately after the previous phase ended. Each phase featured a unique combination of $OFOV \times Set$, so that each participant was exposed to each OFOV level and Set level only once, to prevent carry-over effects. I counterbalanced the order of the $OFOV$s and Sets among the participants.

On the first day, participants signed a consent form and were informed about the experiment procedure, the task, and safety procedures if they experience dizziness or discomfort during the experiment. Finally, participants watched a video guide on how to use the application before proceeding to the beginning with the first phase of the experiment. Each phase included a tutorial stage in which participants familiarized themselves with the interface and the varying OFOV; a training stage in which participants had to memorize the location of the 15 target pairs in the assigned Set; and a retention stage that evaluated how well participants recalled the learned locations 2 days later. The tutorial and training stages were conducted in VR, where participants used the HTC Vive controller to interact with the virtual environment. The retention stage was conducted on the physical counterpart of the virtual training environment and participants used a laser pointer instead of the controller.

Tutorial Stage The tutorial stage consisted of a single memorization task followed by a recall task. The memorization task required participants to learn the location of three target pairs, while the recall task required participants to recall the location of the learned locations. The explanation for the memorization and recall tasks in more detail for the training stage, can be seen below. During the tutorial stage, participants could finish the recall task only after correctly answering all three trials. The targets in the tutorial were the same for all OFOV levels and did not overlap with any of the *Set*s used during the training stage.

Training Stage The training stage consisted of four iterations of the memorization and recall tasks conducted on the same Set of 15 target pairs. Each target pair consisted of two targets, a red and a green square that appeared somewhere on virtual boards surrounding the user (see below for more details on the layout). To avoid confusing participants I ensured that no two squares from different target pairs overlapped over the duration of the experiment. I opted to use target pairs over a single target to ensure that the presented information would not be entirely visible within a small OFOV.

Memorization Task:

Participants initiated the memorization task by pressing onto the main button of the controller. During the memorization task participants were shown the target pairs in random order. Whenever participants confirmed that they memorized the location of the currently shown target pair by pressing on the main button of the controller the next target pair appeared. After participants were shown all 15 target pairs twice, the system automatically progressed to the recall task.

Recall Task:

The recall task evaluated how well participants memorized the location of the targets. For each target pair one of the targets was shown to the participants who had to recall the location of the missing target. Participants could select one of the squares on the boards by pointing at them with the controller. During the recall task a visible ray was cast from the controller that indicated to participants where they are pointing at. Participants confirmed their selection by pressing on the trigger of the controller. Each selection counted as a trial, which was judged as correct if participants selected the correct square on the boards. A visual indicator informed participants if they answered correctly or incorrectly. Participants had to confirm that they saw the feedback by clicking on the main button of the controller. The system then proceeded to the next target pair, which was selected randomly from those not yet evaluated. An in-application timer for the training stage recorded the participants' response times. The training program also recorded head movements of the participants by storing the VR HMD rotation data. The recall task finished when participants completed a trial for each target pair of the Set. If necessary, the system automatically proceeded to the memorization task for the next iteration.

During the recall tasks, I evaluated how well participants memorized the target pairs through a series of trials. In each trial, I randomly selected which of the targets was shown to the participants and asked them to indicate where the other target from this target pair was. The trial was counted as successful, only if the participant selected the correct location on the board. As such, each recall task had a maximum score of 15.

Retention Stage To judge how well participants could retain the learned information, the retention stage featured a transfer task 48 hours after training phase completion. This transfer task was conducted on a physical setup that was in a different room. After participants were seated on the stationary chair in the center of the boards, they were handed a laser pointer that they could use to point at the board, similar to the controller during the recall task. A researcher indicated the location of a target with a laser pointer and the participants were asked to indicate the location of its pair and to verbally confirm the location they point at. The researcher then noted down the selected location and indicated the location for the next target pair, selected in random order by a program. During this retention stage, the researcher recorded the time with a stopwatch. This was repeated until the researcher inquired about the location of all target pairs in the memorized Set. During the retention task participants received no feedback on their performance. After completing the retention task, Participants were asked to fill out a NASA-TLX Questionnaire, marking the completion of a phase.

After finishing phases 1 and 2, participants were directed to the room where the tutorial and training stages were conducted and proceeded with the next phase. If they completed all three phases, I asked them to take part in a short semi-structured interview where they could provide feedback on their experience and memorization techniques they might have used.

The experiment was approved by the institutional review board of our university and took approximately 90 minutes on Days 1, 3, and 5 and less than 20 minutes on Day 7. Participants received a remuneration of JPY 2000 (USD 18).

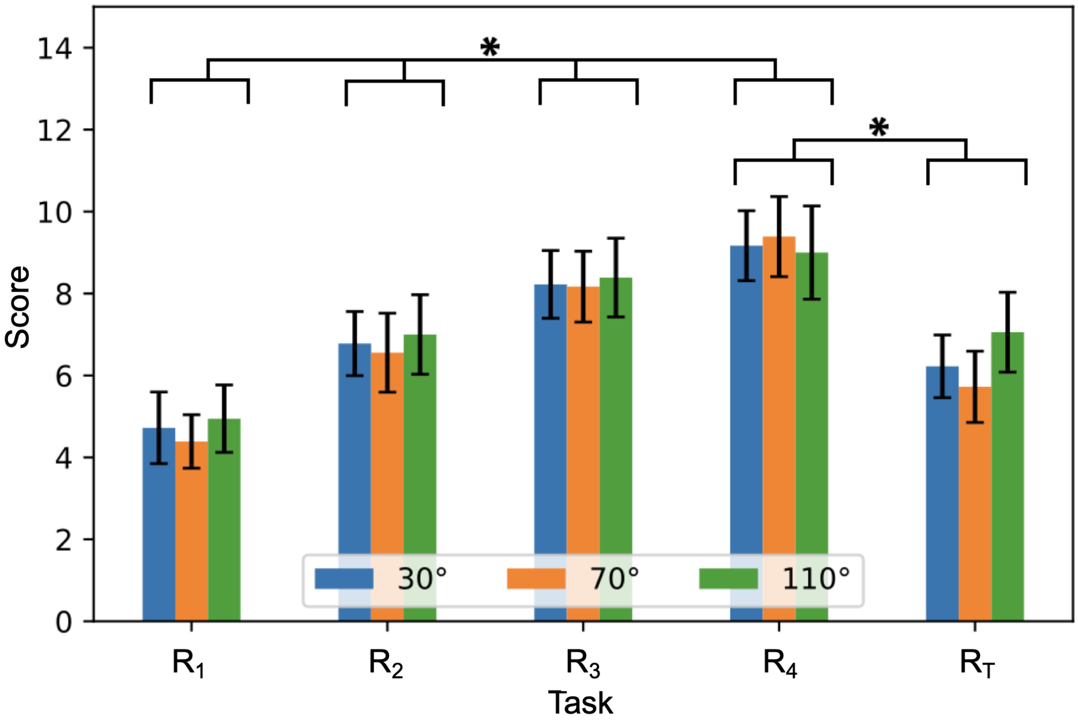


Figure 2.5: Test scores according to *FOV* size and *Task*. Each training score is significantly different from one another. The retention scores are also significantly different from each other. (* p<0.05 ; Error bars represent standard error)

2.4 Results

I compare the objective and subjective performance with the different overlay FOV sizes during the experiment. I determine statistically significant differences using a threshold of 0.05 and apply Bonferroni correction for multiple comparisons.

2.4.1 Test Scores According to Task

I categorize the results of test scores into two: the training scores (considering $R_1 - R_4$) and the retention scores (considering scores from the last recall task in R_4 and the transfer task R_T 2 days after). In total, participants answered 3240 trials for the short-term recall tasks, and 810 trials for the transfer tasks. The summary of the test score results are shown in Fig. 2.5.

Recall Test Scores According to FOV Levene's test did not detect a significant difference between the levels of variance across the recall tasks (F(11,204) =1.278, p ≈0.274 for $R_{1,2,3,4}$). However, a Shapiro-Wilk test showed that the results were not normally distributed p<0.001. Thus I carried out a Friedman's test for all these observations with Bonferroni correction for the adjustment of p values.

Considering the training scores, I did not find any statistically significant difference among the *FOV* sizes, however, there was a statistically significant difference for *Task* (R_1 vs. R_2 vs. R_3 vs. R_4) from the results of Friedman test, $\chi^2(3)$ =46.067, p<0.001. A post-hoc test using Fisher's Least Significant Difference (LSD) showed that all tasks are significantly different from each other. However, a more conservative pairwise Wilcoxon signed-ranks test showed no statistically significant difference between R_3 and R_4. Furthermore, the scores in R_3 were higher than in R_2, which were higher than in R_1. In terms of the training scores, the results reject the hypothesis H_1A.

Retention Test Scores According to FOV Levene's test did not detect a significant difference between the levels of variance across the retention tasks (F(5,102) =2.116, p ≈0.069 for $R_{4,T}$). Like the treatment for the recall test scores, I also carried out a Friedman's test with Bonferroni correction because the distribution of the results were not normal (Shapiro-Wilk: p<0.001).

For the retention task, Friedman's test did not reveal a statistically significant difference between different *FOV*s. Both the LSD and Wilcoxon tests (Z =- 3.427, p<0.001) showed that R_4 had a higher score than R_T. In terms of the retention scores, the results reject the hypothesis H_1B. Overall, the results reject the main hypothesis H_1.

2.4.2 Test Scores According to Proximity of Target Pairs

As the recall and transfer tasks have an equal number of *Adjacent*, *Near*, and *Far* targets, I considered further analysis with *Proximity* as an additional variable. Since the assumption of normality is violated, I used Friedman test for these test scores. To find possible interactions among *FOV*, *Task*, and *Proximity*, I ran an Aligned Rank Transform (ART) procedure [73].

Taking the training scores into consideration, while there was no statistically significant difference in terms of *FOV*, the results showed that *Proximity* had statistically significant difference ($\chi^2(2)$ =29.662, p<0.001), as well as *Task* ($\chi^2(3)$ =46.067, p<0.001). The results from ART showed that there was an interaction between *FOV* and *Proximity* ($p \approx$0.043). After performing contrasts pairwise comparisons, results showed that the scores and FOV at 30° and 70° depended on the *Proximity. Adjacent* target pairs viewed at 70° and *Near* target pairs viewed at 30° produced better scores than their counterparts ($p \approx$0.042). Furthermore, *Near* target pairs viewed at 30° and *Far* target pairs viewed at 70° produced better scores than their counterparts ($p \approx$0.021).

Both Fisher's LSD and Wilcoxon signed-ranks test revealed that at p<0.05, *Adjacent* targets had higher scores than *Far* targets, and the *Adjacent* targets had higher scores than *Near* targets. For Wilcoxon's test with continuity correction, scores from *Adjacent* target pairs were statistically significantly different than scores from *Near* (Z =- 9.392, p<0.001) and *Far* (Z =- 8.347, p<0.001) target pairs. Furthermore, *Far* versus *Near* target pairs had statistically significant difference (Z =- 2.405, $p \approx$0.016.

In terms of retention, the results show a similar statistical significance on the main effects (*Task* and *Proximity*), but did not reveal any interactions among the variables. Using the Friedman test, *Proximity* ($\chi^2(2)$ =28.727, p<0.001) and *Task* ($\chi^2(1)$ =18, p<0.001) had statistically significant difference.

Fisher's LSD test revealed that in the retention scores, with p<0.05, *Adjacent* targets had higher scores than *Far* targets, and the *Adjacent* targets had higher scores than *Near* targets. However, after performing a more conservative Wilcoxon's test, *Far* versus *Near* target pairs had no statistically significant difference. Scores from *Adjacent* target pairs were statistically significantly different than scores from *Near* (Z =- 6.124, p<0.001) and *Far* (Z =- 5.545, p<0.001) target pairs.

I have specified that closer targets would have a higher recall rate. I have to qualify the support and rejection for H_2. The results partially support this hypothesis as *Adjacent* targets yielded higher recall and retention among farther ones. But as *Far* targets had better scores on the short-term recall tasks than *Near* targets, the results do not fully support H_2. The summary of the scores

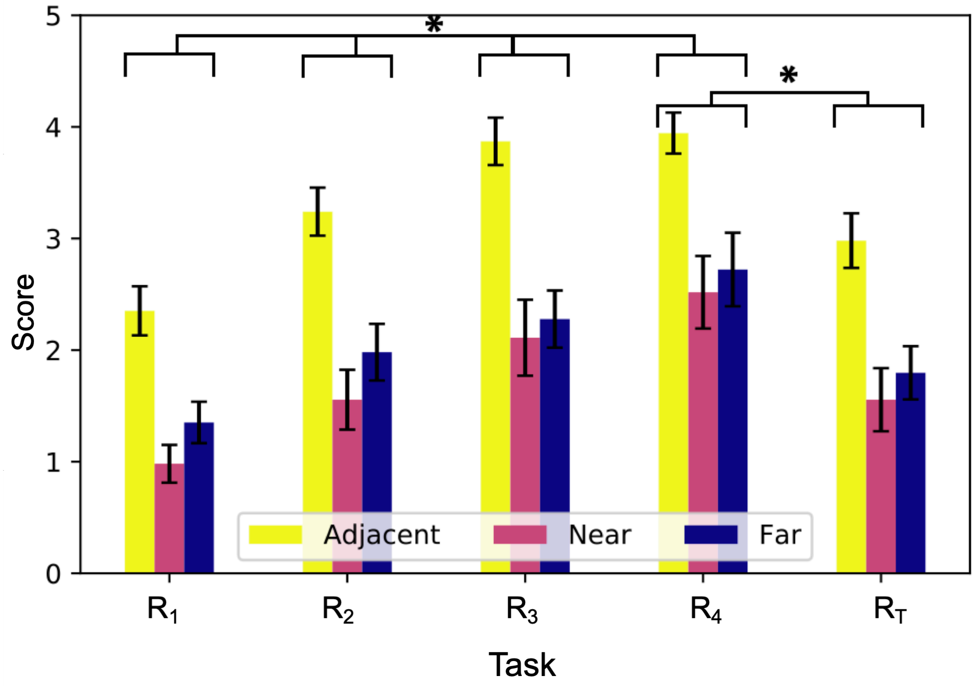


Figure 2.6: Test scores sorted according to *Proximity* and *Task*. Training scores are significantly different from one another. The retention scores are also significantly different from each other as well. (* $p<0.05$; error bars represent standard error).

according to *Proximity* is shown on Fig. 2.6 and 2.7.

2.4.3 Task Completion Times and Head Rotation

The results for task completion times and head rotation did not violate the normality and sphericity assumptions. I thus analyzed them using an ANOVA with a post-hoc Tukey test.

A 2-way repeated measures ANOVA found no statistically significant difference between FOVs, and response times for the short-term recall tasks and the transfer task (Fig. 2.8). It also did not reveal an interaction effect between FOV and tasks. This result thus rejects the hypothesis H_3.

A 2-way repeated measures ANOVA also found no statistically significant difference among memorization task completion times, whether by *FOV* or by *Task*. This is also true when looking at the entire duration of one day of training.

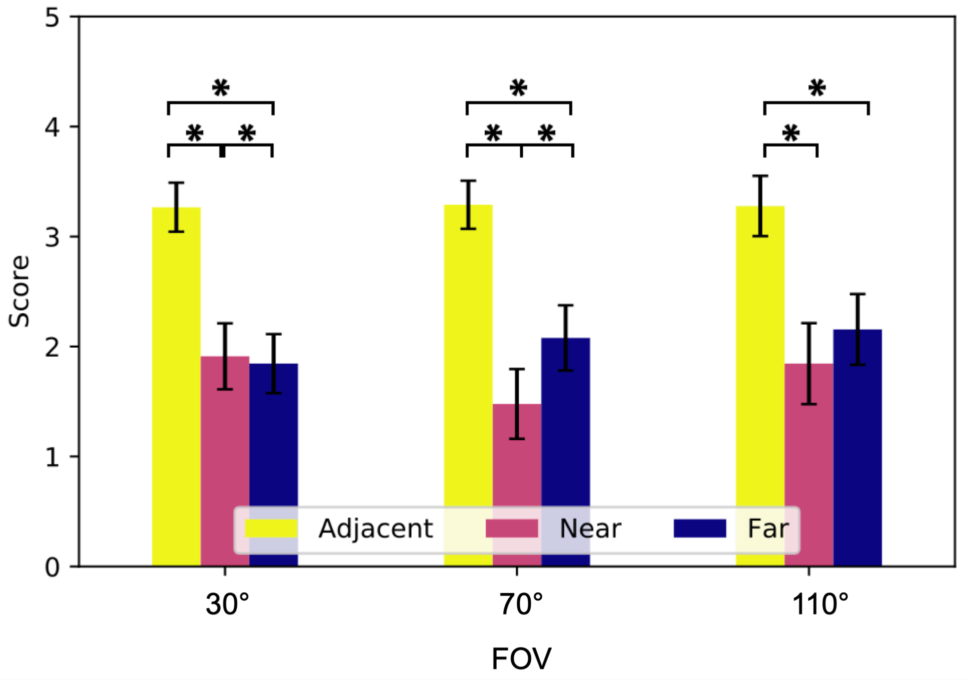


Figure 2.7: Test scores sorted according to *Proximity* and *FOV*.(* $p<0.05$; error bars represent standard error).

I presented the distribution of head rotation data as a heatmap plot (Fig. 2.9). The top-down view of the experiment shows the relevant y-axis (up added) around which the participants' heads rotated. To analyze the head rotation data from the HMD, I processed the movement as angular velocities around the y-axis, where participants rotate to view different panels during the training stages. I can see that participants rotated towards the leftmost and rightmost panels more as the FOV size got smaller. This indicates that they had to rotate their heads to complete the view of the target pairs, especially for the 30° FOV. On the other hand, the heatmap indicates that most of the time, the viewport of the HMD remained centered and participants did not rotate their head much in the 110° FOV since the larger FOV size can show more objects.

I looked at the average shift in head rotation around the y-axis, by taking the sum of all the values of the rotations over the duration of one session. I designated the starting time of the training as the first click of the controller to start the session, and the ending time of the training as the last controller click

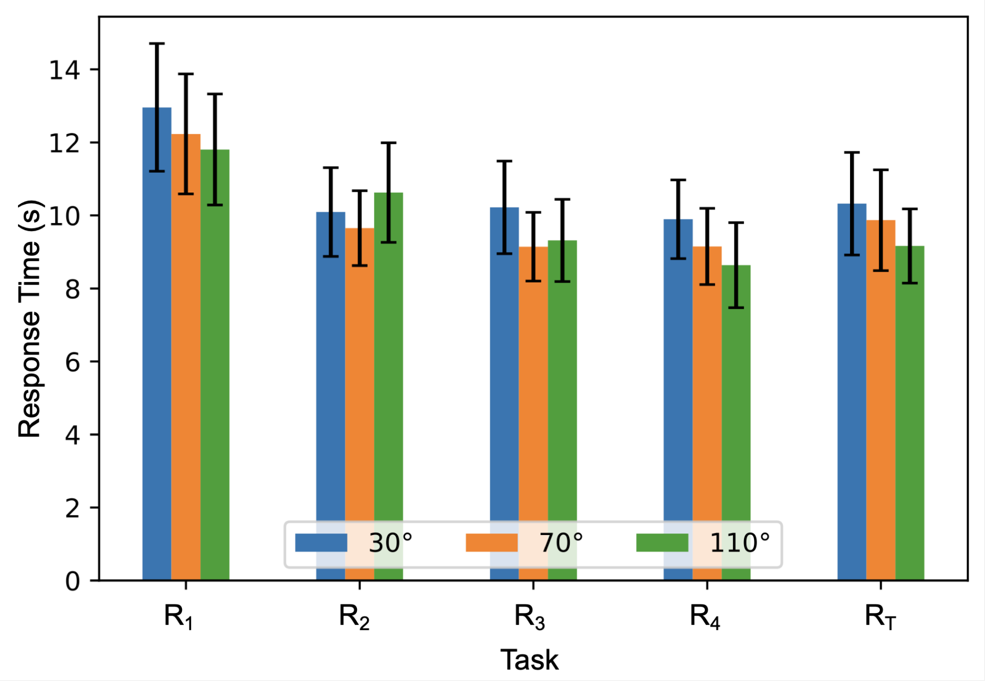


Figure 2.8: Average response times while performing the answering tasks according to *FOV* size and *Task*. Error bars represent standard error.

to answer the last item of the R_4 recall task. The difference between the starting and ending time is the duration of the session.

Before the analysis, I have removed three trials as outliers from the data. Then, I performed a repeated measures ANOVA, reporting a statistically significant difference in terms of FOV ($F(2{,}48) = 8.045$, $p \approx 0.001$). A post-hoc Tukey test revealed that velocity values of 110° were smaller than those of 30° and 70°, as both pairwise comparisons were statistically significant with adjusted p values of $p \approx 0.004$ and $p \approx 0.002$, respectively. The results thus only partially supported the hypothesis in H_4. Showing that although the FOV did not significantly affect the training time, participants exerted more effort rotating their view while viewing smaller FOV sizes.

2.4.4 Preferences and Evaluation from Participants

After the experiment finished, I asked participants different aspects of the training beyond the link between FOV and spatial memorization to get subjective feedback

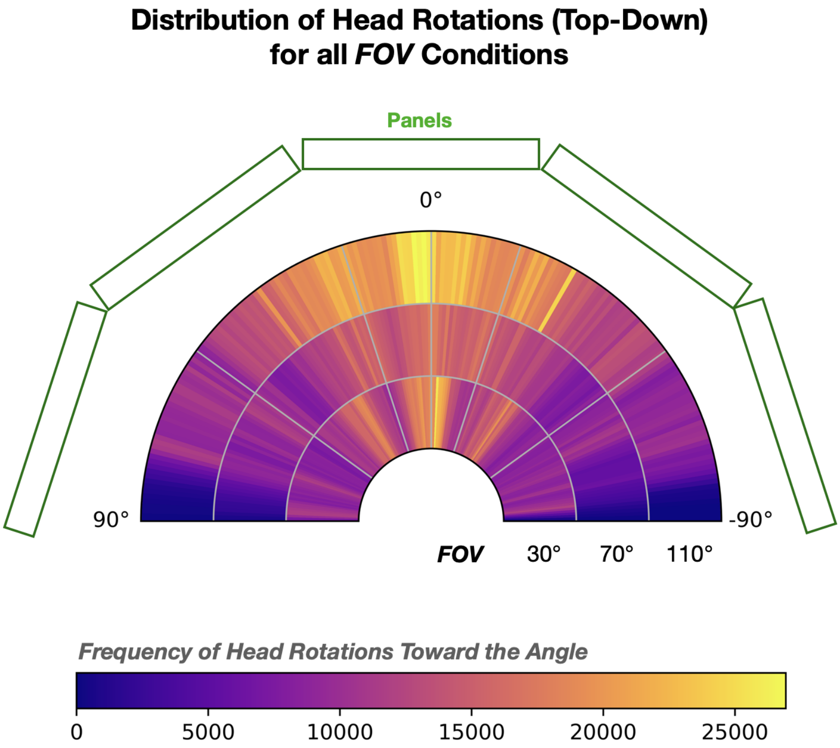


Figure 2.9: Visualizing the head rotations from the top-down view for the three different levels of *FOV*, with each level separated by the radius grid lines.

about how they approached the task. The intrinsic motivation of the participants vary. Three participants directly stated that they did not exert too much effort to get correct answers, guessing or skipping items they did not memorize. One participant directly expressed seriousness in getting all the answers correct. Half of the participants additionally expressed being overwhelmed about the number of items to memorize. Six participants also noted that memorization became easier as days progressed.

Since this was a task about memorizing spaces, it was important to ask about their memorization techniques. Most participants answered that they relied on the 2D grid structure of the panels to memorize. Different participants had different ways of dividing the memorizing space. When I asked participants about

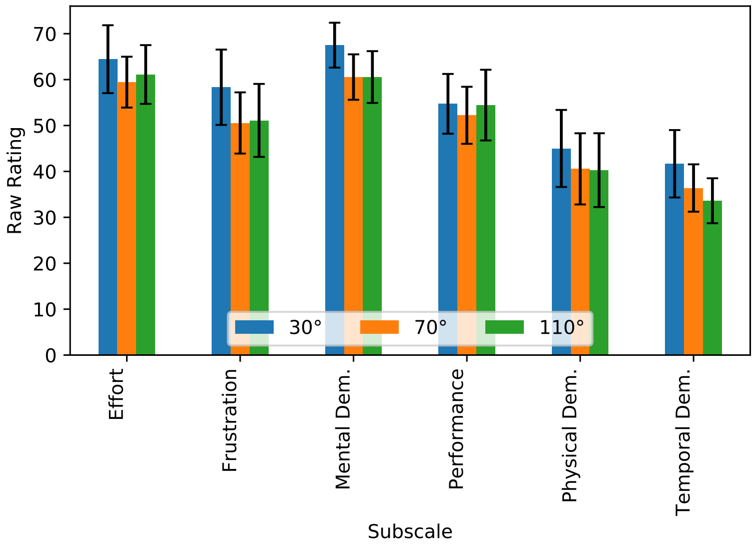


Figure 2.10: Results from the raw NASA-TLX questionnaire. Error bars represent standard error.

memorization strategies, I found that they used a combination of similar approaches, which I classified into three. First, seven participants created a matrix/coordinate system for numerical associations of the locations of the targets. Eight participants broke down the panels into components or regions (e.g., gray line makes a top and bottom block, thus treating the space as 10 boards of 4x3 squares). Finally, nine participants imagined a target pair as part of a shape or a pattern.

I also asked participants to rate the cognitive load of the tasks for each day they performed in the user study, via a raw NASA-TLX questionnaire (Fig. 2.10). Using Friedman test, there was no statistically significant difference.

I wanted to know if participants noticed and relied on the wedge guides when they could not find target pairs. According to participants, they all noticed the wedge guides. I also asked participants on how reliant they were with the wedge guides on a scale of 1 (“I did not care”) to 10 (“I was using it frequently.”), participants reported an average of 7.5.

2.5 Discussion

According to results of the experiment, I was not able to observe a statistically significant difference among the chosen size of the overlay FOV of the HMD.

The task takes into account how fast and accurately the participants were able to retrieve spatial information of a pair of locations in 3D space. In this section, I match the objective results and participant reports with ideas in spatial memorization and perception that may have played a role in the study.

2.5.1 Test Scores According to FOV and Task

I was not able to observe a statistically significant effect of the FOV size on the scores of the spatial memorization test, whether in short-term recall tasks or the transfer task. As expected, the results also showed that after a few repetitions of training, the performance in the recall tasks improved, but not all of the memorized items were transferred into the real world. To explain the results in different perspectives, I refer to two known processes about memory that differ in terms of the amount of physical and mental effort exerted: automatic and effortful [29].

First, there is automatic registration of memory, based on the visuospatial sketchpad idea from the working memory model [6]. In this experiment, I assumed that this automatic memorization benefited from a large FOV where participants could effortlessly see both target pairs, regardless of proximity of target pairs. However, the results did not reflect any statistically significant change among FOV sizes. Participants reported that they could easily memorize the location of some targets, recalling them correctly in the first tests. These targets usually could be easily associated with features of the environment, e.g., close proximity to the edge. Here, the edge likely functioned as a landmark.

Second, adding intent to the process of memorizing spaces produces an effortful training, improving what is being registered automatically [17]. In the case of memory palaces [40, 45], participants had to exert effort and change location to improve memorization, whether by locomotion physically or by scene movement virtually. This intent may also manifest as “muscle memory”, or associating memorizing a location with a specific motor movement, like the amount of necessary head rotation being linked to the distance between targets. Montello et al. [51] explained the role of the body in the spatial memorization process. They highlighted how humans update knowledge of their locations based on perceived body speed and direction. If participants would rely on this muscle memory, this assumes that users exerted extra effort, which might be the reason I detected no significant difference among FOV sizes according to the retention scores and their response times. Based on the results of the experiment, inducing effort through reducing FOV may not be relevant to the performance.

The linking of perception of the training environment to actual performance still needs further clarification. On one hand, the participants have reported different techniques and approaches to their memorization process. On the other hand, the inherent motivation to actually memorize the sets were different for each participant. Some participants were very eager to get all of the answers correct in as few tries as possible, while others decided to skip hard to memorize items in favor of the ones that they have already memorized. While these techniques and motivations can somehow connect to the fact that ease of the memorization process can be aided with a certain size of the FOV, there still needs to be more direct evidence for further examination.

Finally, results showed that scores had a significant difference according to *Task*, not *FOV*. Participants had more chance to immediately improve scores on short-term recall tests, as these happen as soon as the memorization task finishes. Hence, as expected, a learning effect was observed on the training stage, and a performance deterioration on the retention stage. As expected, the task was challenging, with participants correctly recalling only 9 of the 15 pairs in R_4.

2.5.2 Test Scores According to Proximity of Target Pairs

Proximity of the targets also had a role in the memorization process. Looking at the results of all tests and the consistency of participants' observations after the whole study, correctly-answered target pairs were closely linked with proximity: either the target pairs are close to each other that they fit the overlay FOV (which were confirmed to be the *Adjacent* target pairs), or they are on the extreme opposite ends of the whole layout (which were the *Far* target pairs). From these observations, I emphasize the importance of spacing and grouping targets together.

The anticipation of objects in the periphery has been discussed by Ragan et al. [58] as an affordance of a large FOV. The results showed that for a large FOV like 110°, farther targets (*Near* and *Far*) produced similar outputs. This may indicate that at a certain threshold value of FOV size, users may regard the periphery with the same level of anticipation. Furthermore, in the case of *Near* target pairs, these positions represent locations found near the immediate periphery of the overlay area when a person does not perform head rotations. This combination of a lack of head rotation, a shift away from the visual center, and an even restricted view brought about by small FOV sizes may have contributed to the participants forgetting such pairs.

2.5.3 Task Completion Times and Head Rotation

As mentioned, some participants only focused on target pairs that they were sure they memorized. For items that they could not memorize, they either guessed or opted to skip by selecting a random location. On the other hand, those who had a strong intrinsic motivation to memorize everything and to have a perfect score, with an intention to answer carefully. Thus, response times during recall tasks have varied.

Furthermore, the distribution of the head rotations along the y-axis map well with the *Proximity* classification of the target pairs on the corresponding panel. The *Adjacent* target pairs are easily viewed with larger FOV sizes because there was no truncation, while *Far* targets are more difficult to view in one frame when the FOV sizes are smaller.

Looking at the distribution of rotations from the top-down view (Fig. 2.9) and the results of the rotation data analysis, the 110° FOV had the least rotation. This meant that most of the time, participants did not have to shift their attention somewhere else. Given that they had the maximum FOV size at that moment, they could accommodate the greatest amount of visual information. The 30° FOV demonstrated the opposite, where participants had to rotate themselves to the left and the right more than the other two *FOV* levels. This implies that participants had to exert more effort to visually inspect the scenario and locate the target pairs, aside from the actual memorization itself.

The statistically significant difference in the average velocity revealed that the physical performance of the task is affected by the FOV. Although, the more appropriate observation was that with a large FOV of 110°, participants did not have to rotate much more than how they did with the other two FOV sizes. Aside from this, given that the velocity values for 30° and 70° did not have a statistically significant difference, it would be interesting to investigate the threshold FOV size where the rotation speed values would change significantly.

2.5.4 Preferences and Evaluation from Participants

The tendency to remember locations with respect to positions within the grid has been explored [46], citing the observation about a worse performance for content memory in the presence of grid lines. This supports the participants' claims of ignoring the content of the target pairs (the colors red or green) as the training sessions progressed.

In terms of the NASA-TLX scores, Friedman's ANOVA may have been too restrictive to produce statistically significant differences. However, the difference could potentially become significant with more participants. The high evaluation score and the overall awareness of the presence of the wedge guides indicate that these visualizations of off-screen virtual objects are beneficial for participants wearing OST-HMDs.

2.5.5 Visual Characteristics of the Training Environment

When designing the experiment, the task and the environment were major considerations. To keep the results consistent, certain parameters have already been set in place.

The view of the training environment in the experiment that emulates the view of the person wearing an actual OST-HMD may play a role. Since the scene is a simulation of an AR environment, the view consists of the overlay FOV, where the virtual objects can be seen occluding the real world on the background, and the peripheral FOV where only the real world can be seen. In this scenario, the application still presents a constant view of the real world in its fullest extent, which does not always match the size of the overlay FOV. Most studies [2, 14, 39] limited the overall FOV, thus also limiting the view of the real world, leaving a blacked out space outside of the viewable area. However, this experiment only adjusted the overlay FOV. The availability of the peripheral FOV as the source of visual information may have still kept a frame of reference for participants to use. Thus, regardless of the FOV, participants still had access to crucial landmark information provided by the real world to mentally construct a spatial map.

Aside from the distinction between the overlay view and the real-world view mentioned earlier, saliency of the virtual objects against the environment is also important, as identifying such elements that are distinct from a regular perception of the environment would assist participants in their visual judgment of the scenario. Hassan et al. [30] have made this distinction important by pairing up contrast levels with specific minimum FOV sizes to achieve an efficient navigation. Trepkowski et al. [70] has also pointed out how information density, when linked closely to the small FOV sizes of HMDs, could affect search performance.

Given these studies, the environment went hand in hand with information on the display. The main idea behind using only the monochromatic background (i.e., black, white, gray) was to ensure that the rest of the colors can be designated as the colors of the training objects. In this user study, red and green were the target objects, while blue served as the laser pointer color during the training stages. Unlike the more detailed virtual environment [40, 45], this work eliminated extraneous details which might be considered important virtual landmarks.

This training environment with minimal color choices helped with minimizing the mental load of memorizing. Participants did not mind the actual positions of the colors, and instead focused on the spatial relationship. Furthermore, unlike studies on virtual memory palace applications, the stimuli for this experiment were strictly visual, without any additional complexities brought about by faces [40] or imageability of words [45].

Having to memorize a pair of two stimuli is also just one particular kind of memorization task. This type of task has been selected since the capacity to memorize them is low (i.e., only two targets with simple colors), in a fixed location (i.e., the target locations and the panels do not move), and the order or sequence does not matter (i.e., they must be memorized together as the test randomizes the missing target). While these properties have been simplified for this experiment, other tasks like memorizing assembly of objects require step-by-step procedures which include an additional layer of episodic memory, and not just semantic memory.

2.5.6 Experiment Limitations

Primarily, the maximum FOV of the device as a restriction also became the limitation of the experiment itself. Deploying the experiment in an HMD (whether as an immersive HMD or OST-HMD) that would accommodate FOV sizes from very large (preferably as wide as the human vision) to small would be ideal. I also investigated only one spatial memorization task. There may be other tasks that require skill transfer, spatial or otherwise, for which FOV size could play an important role.

The design of the environment is also something that should be noted. In real and actual work spaces, many environment changes happen, like lighting or potential distractions, and workers may move around instead of just sitting down. This experiment only featured a static scene for the training application, as participants experience the training in a stationary position. As the experiment restricted the movements to egocentric rotations and head movements, the relationship between FOV and dynamic factors were not observed (e.g., the optic flow of the scene, potential visual interruptions, or extraneous visual stimuli). Allowing movement would be an interesting future work.

The experiment also presented a very unique memorization task. The recall test was randomized instead of a free recall or serial tests, and I have also presented the order of target pairs in a randomized manner during the training stages. While this experiment may have eliminated order effects in the memorization process, a different procedure for assessment may yield different results and insights.

I assume that by mixing either episodic or semantic memory to a spatial memorization task will demand more cognitive resources, and thus affect the overall memorization performance of participants. For example, when participants would additionally need to follow a specific sequence of red and green targets along with their respective locations, participants may solely rely on the sequence instead of locations. On the other hand, prior semantic memory of icons, characters, or words might interfere with spatial memory, as they carry additional meaning over stimuli that only differ in two colors.

Another potential limitation of this study is the timing of evaluations outside the recall tasks. I asked participants to answer questionnaires like NASA-TLX the moment before another set of recall tasks were administered. The significantly long time that elapsed between the end of the last recall task and the moment I have given them the questionnaire may have affected the results of these subjective evaluations.

Furthermore, keeping the participants on a consistent schedule of coming back every 48 hours posed a challenge. The 7-day schedule meant that participants needed to come back on the weekend to finish the whole user study.

Finally, intrinsic motivation to perform the training task can also be improved. If the goal of all participants is to perfect the score, they can equally pay importance and attention to all the objects they should memorize. Other dynamic interactions that would encourage the learning of content should be explored, rather than mere visual search or inspection of the targets in a stationary position.

2.6 Conclusions and Future Work

In sum, I presented an experiment that investigates the effects of the overlay FOV size in HMDs on task performance, particularly, on the skill of memorizing spatial maps. To provide a demonstrative example, this study focused on the spatial relationships between two virtual objects. However, results suggest that the selection of FOV sizes did not significantly contribute to any negative or positive task performance. Furthermore, it did not affect any aspect of the memorization, from the correct responses, to the task efficiency indicated by the response time, to participants' preferences. However, the large FOV (110°) showed the least amount of head rotations still indicates an ergonomic advantage when performing repeated training.

The main recommendation is to extend this study with more participants and modify the parameters of the study. The post-experiment interviews revealed that the performance of the task is still influenced by many other factors like confidence or intrinsic motivation, and not directly by the perceptual issues due to limited FOV. Thus, this experiment provides a counterpoint to the criticism that the FOV size in HMDs directly affects task performance. For companies and organizations that are deploying or would like to deploy AR-based modules to trainees, choosing an OST-HMD may not matter in terms of the actual task performance, but the ergonomic advantage of having a larger FOV size should still be taken into consideration. It is also important to analyze what kind of memorization techniques the trainees are familiar with, so that a specific category of FOV size accompanies such technique.

Even when previous work [30, 70] have identified critical FOV angles for an efficient spatial task performance, they only accounted for minimum FOV angles with a certain visual aspect. This experiment did not have this correspondence, as all the conditions have the same contrast levels or information density. For future work, another suggestion is to look at better FOV size thresholds and critical points, or re-frame the experiment with a relevant visual aspect. Identifying the minimum FOV size that will not affect spatial memorization will just be as important as identifying the starting large FOV size where the performance will stay the same even if the FOV size increases.

I also suggest exploring other training scenarios where FOV size is crucial.

This study defined the “complete view” as the appearance of two virtual objects (or possibly more) on the same view, which meant that the farther these objects are from each other, the less likely they can be accommodated by a narrow FOV. It may be important to explore the consideration of “complete view” as large-scale virtual objects or information that is larger than the FOV allows.

Finally, as this experiment was a simulation of an OST-HMD, and visibility of the real world may affect the results it would be important to repeat this experiment on an OST-HMD capable of augmenting a large FOV.

3. Influence of Perception-Based Augmented Reality on Driving Speed Perception

When AR superimposes visual information onto the real world, the resulting view may influence users' perception. However, what the user expects to experience sometimes does not match what the AR system renders onto the scene, causing a misperception. In this chapter, I enumerate the differences between the conventional physically-based AR and perception-based AR which considers the limitations of human perception. Furthermore, I argue that by observing the perceptual process of speed, the AR application can use the inverse of this process when displaying a scene with virtual objects that are perceptually consistent with their intended target appearance. To explore and evaluate, I conducted an experiment that looks into the influence of information density in OST-HMDs on driving speed perception. The information or stimuli is presented as a radial vection pattern, to be shown against a nighttime driving scenario that lacks many visual cues. The general results show that displaying sparse and medium densities (32, 64 moving stimuli) would already affect the driving speed perception. However, individual task performance indicates differences in preferences and receptiveness.

3.1 Introduction

Perceptual issues in augmented reality (AR) presented open challenges for research and development of better applications and systems that suit or even enhance the limited human vision. The design of AR systems must consider limitations of human visual perception and various visual complexities of the real world.

To resolve such perceptual issues, the perspective of developing and evaluating AR experiences must shift from a physically-based one to a perception-based one. The use of perception-based AR (Fig. 3.1), differs from physically-based AR in terms of approach, process, evaluation, and end goal.

The motivation for this study is for AR researchers and developers to consider this perception-based paradigm in the design of new AR applications. This approach also uses the results from studies that explored the effects of different

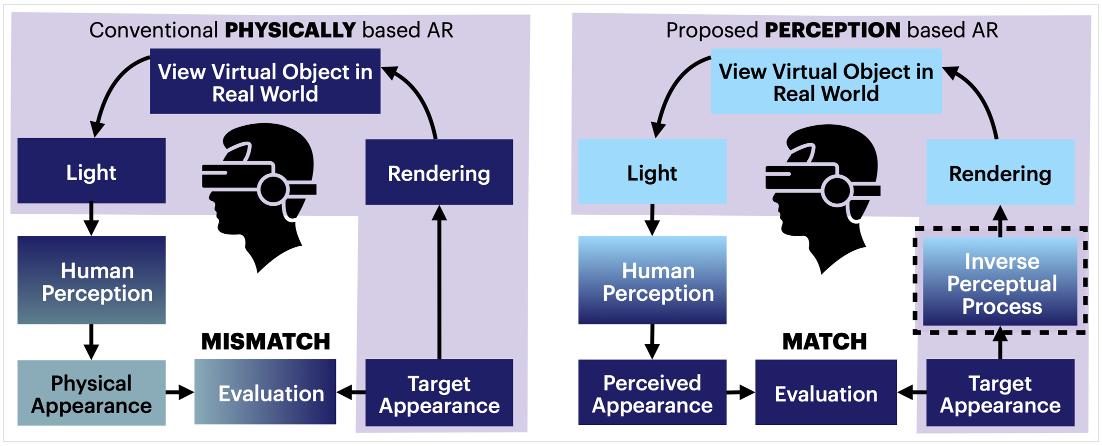


Figure 3.1: Physically-Based and Perception-Based Augmented Reality differ in terms of the focus of evaluation and a key pre-rendering process.

factors in AR systems on human perception.

I describe how other related work have already used this paradigm of perception-based rendering, through experiments and applications. Then, I present an application of this paradigm through an experiment that investigates speed perception while wearing an OST-HMD in a driving scenario.

3.1.1 Physically-Based and Perception-Based AR

In a typical visual rendering pipeline in AR, the display device would render the scene according to the device capabilities of the system. The virtual objects are presented as light superimposed onto the real world. Then, the user of the system would perceive both light and the surrounding real world details.

For physically-based AR, the important aspect of evaluation is the match between the target appearance and the resulting output just before the user perceives the scene (i.e., in terms of visual aspect, this is light after rendering). The benefits of these kinds of evaluations can be seen through applications that involve accurate measurements, like in surgical operations [75] or motion graphing [12]. In these cases, users can experience the visual aspects of objects that may be difficult to perceive without the AR system (e.g., true and accurate values of size, speed, order of appearance from human view).

Physically-based AR considers only a straightforward rendering of the scene, without deliberate consideration of how the user will perceive such a scene. When the light enters the user's eyes, the appearance of the virtual objects may change.

Furthermore, physically-based AR only focuses on the evaluation that follows the laws and phenomena of the real, physical world as the basis of visual fidelity of computer-generated content. This physical evaluation is different from the human-based evaluation of perceived appearance. This causes a mismatch between the originally intended appearance with what the users actually perceive.

In conventional physically-based AR scenarios, the rendering relied on the physical properties of the real world. Visual complexities from the real world such as contrast or artifacts may interfere with the rendered content, and the resulting appearance may be perceptually distorted. However, in perception-based AR, the pipeline has an additional *inverse perceptual process*, which anticipates any perceptual change that may occur even with the presence of interference from the real world. This inversion borrows from previous research on the processes of human perception and opposes or modifies them to create a match between the perceived appearance and the target appearance.

As a caveat for this chapter, I will not immediately demonstrate this inverse perceptual process in the experiment. For this to happen, I need to observe first the perceptual process: how will users with OST-HMDs perform when an AR pattern is shown against a driving scenario in the real world? As such, my approach is to design a rudimentary AR pattern which can be developed and optimized, so that I can learn more about this perceptual process. Once done, the future steps for this research is to formulate the function of the speed perception process and its inverse, and to figure out how to correctly augment this inversion.

3.1.2 Augmented Reality as Drivers' Display

Head-up displays (HUD) offer a user interface that can provide visual information to drivers within the forward FOV. A variety of displays have been used as driving HUDs like projectors [69] or small dashboard interfaces. The HUD can only provide a small area within the dashboard to view the AR content, hence, the attention and focus of the driver shifts back and forth, from the road to the display. Mobile and handheld AR devices such as smartphones do not work in this scenario unless fixed in a position on the dashboard.

Given the drawbacks of HUDs, OST-HMDs like the Microsoft HoloLens 2 offer a number of advantages that can be suitable for this experiment. First, drivers can view a near-eye and egocentric version of the HUD through OST-HMDs, which reduces the shifting of focus. Second, the untethered and compact design of these displays is conveniently placed as close as possible within the FOV of the driver, while still delivering a good quality of AR content.

Current research still investigates the role of FOV of OST-HMDs in the effective delivery of AR content. However, I also see its restriction as an important factor to consider when placing the augmentations to influence driving speed. The application I have developed shows different information densities of a certain vection pattern, rendered as holographic content in a Microsoft HoloLens 2. The behavior of the animation of AR content is consistent with a constant speed over the course of the driving scenario. The goal of the stimuli is to make the OST-HMD user associate the constant speed pattern with a specific target speed, as the stimuli are presented alongside the driving environment.

3.1.3 Research Contributions

The contributions of this chapter are as follows:

- to conduct an empirical study and gather subjective feedback about the relationship between information density in OST-HMDs and the tendency to incorrectly perceive speed (using driving as a scenario),
- to provide insights about the influence of augmented content on the way a user perceives and carries out a task, highlighting the trade-off between seeing virtual objects as a distraction or as a useful stimuli against the view of the real world in an OST-HMD setup,
- to highlight observations made on speed perception and human perceptual process of speed that could have implications on the design of speed stimuli that counteracts a poorly perceived real-world motion (e.g., closing the gap between actual and perceived speed), and
- to present findings relevant to the discussion of whether or not such animated patterns in AR would help users perceive speed correctly.

3.2 Related Work

Perception-based AR relies on the results of experiments that explore how humans perceive an environment in terms of different visual factors, e.g., color or depth cues in AR. Exploiting these factors can already address any misperception before the rendering, and thus create an AR experience that can accommodate the human visual system.

3.2.1 Examples of Perception-Based AR

The color of objects and the environment is a fundamental visual aspect in human perception. For example, a spatial AR system using a projector-camera combination induced a visual illusion of color constancy to broaden the range of perceptually presentable colors of projectors [1]. The manipulation of the colors of the foreground target object and the background environment serves as the inverse perceptual process.

In Fig. 3.2, the original appearance of the image is difficult to transform into the target appearance by merely projecting simple light onto the whole image. The target color(s) of the foreground object(s) and the corresponding results from the naive projection are the same upon evaluation of their physical properties (i.e., the same color values). However, they look different when a human perceives them.

Humans can recognize the colors of objects invariant to the light source through color constancy. This prior knowledge about human perception can be inverted and exploited by surrounding the target object(s) with a different kind of light. The area of the target object(s) in the image is isolated, and a different light source will also be projected onto it. The resulting image using this technique would have a perceptual appearance that is similar to the target image.

Depth cues can also be used in an inverse perceptual process. Various visual factors in rendering can also affect the perception of depth, and consequently, distance judgments. For Diaz et al. [16], there are trade-offs across virtual object designs like the appearance of shadows, shading models, or dimensionality of the virtual object. They outlined a detailed experiment setup and evaluation procedure that compares the perceived distance judgments of participants to the

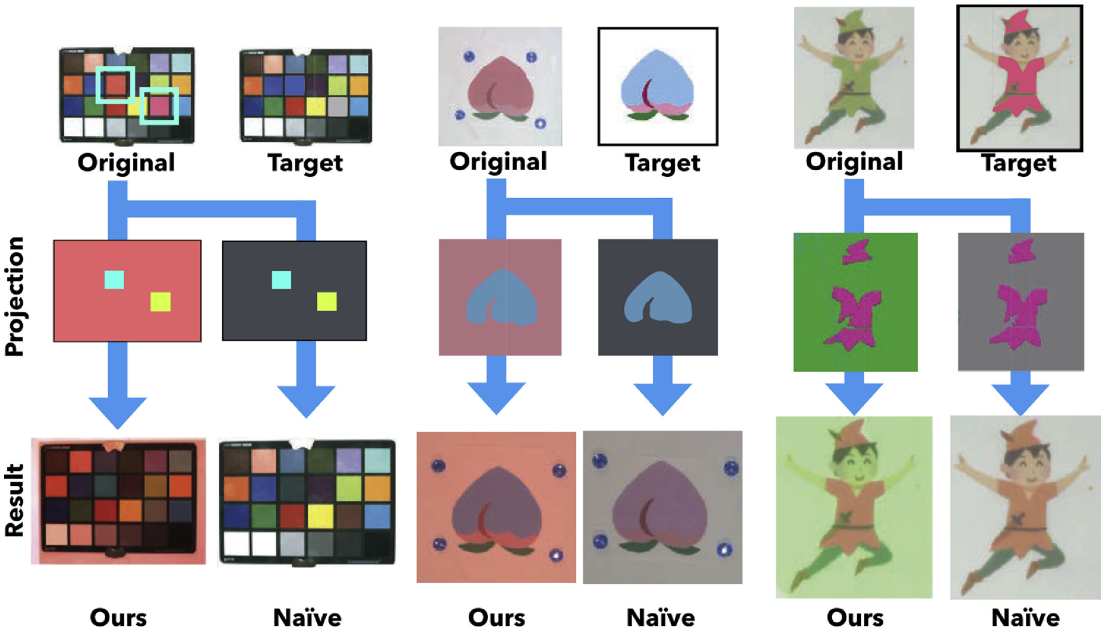


Figure 3.2: The manipulation of foreground and background colors by Akiyama et al. [1] as an example of an inverse perceptual process.

physically-based measurements. Results showed how even a simple cast shadow already improves the participants' distance estimation of the virtual objects.

When these observations about human depth perception is used, systems that automate depth enhancement [10] or introduce auxiliary AR content [42] would be able to invert the perceptual process, and render depth cues in such a way that users can correctly judge depth and shape of the rendered graphics.

3.2.2 Speed Perception in Driving

Humans perceive stimuli which are usually in motion. In this experiment, I focus on human speed perception, or the ability to estimate one's own motion or the motion of other stimuli. This process is integral in the development and training of higher order skills like obstacle avoidance or target surveillance. And like other cognitive processes, many visual factors affect speed perception.

Pretto et al. [55], examined how the field of view (FOV) affected users' speed judgements. As they narrowed the FOV, they found that the lack of visual information in the center or in the periphery led to an incorrect estimation of speed.

Lidestam et al. also observed this relationship between FOV and speed perception in their experiment [48]. Their results showed that driving speed is reduced when the number of monitors surrounding the FOV of the user is increased. Other factors that affected driving speed were rich motion-flow cues and virtual road markings.

Contrast of the environment affecting speed perception have been demonstrated in a driving experiment about the loss of visibility through fog [65]. Their experiment procedure involved the adjustment of an accelerator and a brake to control the speed in a driving simulator. The participants must match their speed with the target speed, while different conditions appear on screen (e.g., clear, misty, foggy). Another experiment by Pretto et al. [54] further explored the fog problem by using fog and anti-fog conditions, aside from the clear baseline condition. Then, they asked participants to compare the speeds of the two.

Aside from identifying the visual factors, studies of AR systems in driving involve the design of AR content which can help drivers recognize velocity or reduce motion sickness. The results from their experiments inform this perception-based AR system to influence the speed perception of a driver.

Sawabe et al. experimented with vection-based designs with the aim of reducing motion sickness while seated on a self-driving car [63]. They presented dots around the scene that moves radially outward, mimicking the optic flow of the driving scenario. Toui et al. developed a projector system that displays lines and shapes onto the windshield for drivers to recognize their current velocity [69]. These designs resize with distance and follow the curves of the road. For driving speed adaptation affected by FOV and virtual road markings, Lidestam et al. also suggested that a driving interface should be non-iconic, naturalistic, and intuitive [48].

The type of presentation also influences the speed at which drivers could react to visual stimuli. Merenda et al. prepared four different types of AR HUD interfaces, which vary between the frame of reference (screen-fixed interface versus world-relative conformal AR) and the behavior of the interface elements (static versus animated) [50]. Their results showed how animated design can assist goal-directed navigation, but may cost additional response time and distance.

3.2.3 Summary of Related Work

To determine if AR-based stimuli would influence the speed perception of a background environment, it is important to examine the properties of the stimuli presented, the nature of the environment, the display used, and the task objective that should involve a judgment in speed perception. As a summary, Table 3.1 shows the differences in components of the related work covered in this section.

Table 3.1: Positioning of this Speed Perception Experiment Among Related Work

Ref	Displays	Task Objective	Stimuli Type(s)	Stimuli AR?	Subjects (#)	Treatment
[48]	Monitors in Hemisphere	Drive at Target Speeds (40, 80, 120 km/h)	Virtual Road Markings	Not AR	62	within
[69]	Projection onto Windshield Screen	Recognize Velocity	World-Dependent Line Patterns	AR	various	mixed
[38]	Projection onto Windshield Screen	Detect Stimuli	Static, Dynamic, Blinking	Not AR	12	between
[54]	Panoramic Quarter Sphere Screen	Estimate Target Speeds (40, 60, 90 kph)	Varying Fog Conditions	Not AR	0	within
[63]	Video See-Through HMD	Experience Driving Motion	Moving Dots as Vection Patterns	AR	9	between
[55]	Panoramic Quarter Sphere Screen	Estimate / Judge Speed	Random Moving Dots	Not AR	10	within
[50]	Augmented Reality HUD	Identify Pedestrians, Locate Parking Space	Arrows and Pedestrian Icons	AR	24	within
[66]	Panoramic Quarter Sphere Screen	Estimate / Judge Speed	Dots and Spheres	Not AR	8	within
This	**Microsoft HoloLens (MR) + 4K monitors**	**Drive at Target Speed (80 km/h)**	**Moving Line Segments**	**AR**	**24**	**within**

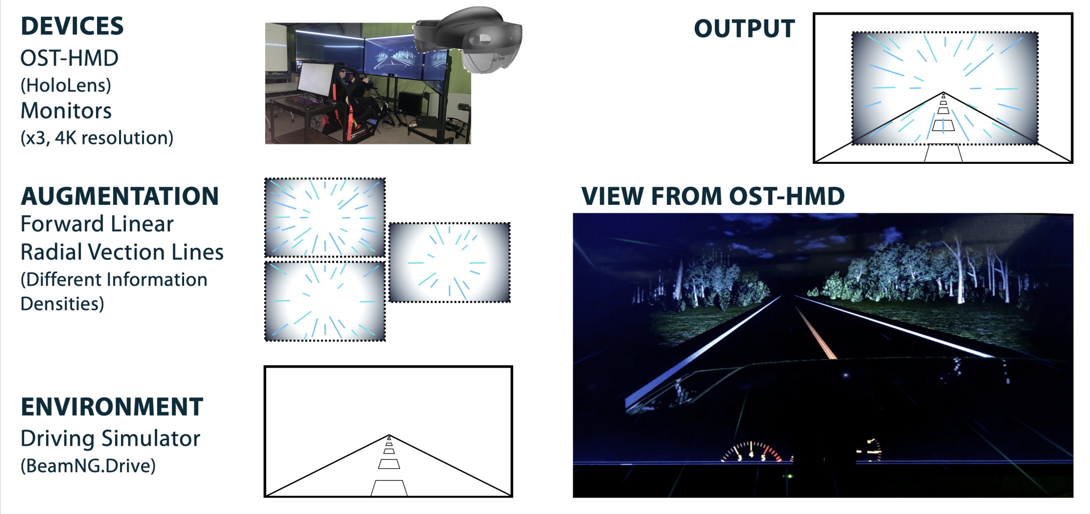


Figure 3.3: Components of the perception-based AR system that aims to influence the speed perception of a driver in the driving simulator, with the AR content as stimuli viewed through an OST-HMD.

3.3 Experiment

A perception-based AR approach to the speed misperception problem involves identifying the properties of the environment, the display, and the augmentation. I present the system for this experiment (Fig. 3.3) that modifies various properties of the visual stimuli rendered in AR to potentially counteract the problems of the driving environment. The aim of this application is to influence the driving speed by using AR content.

3.3.1 Driving Task

The main task for participants is to begin accelerating the car to a specific speed, and keep the same speed throughout the rest of the trial. I selected 80 km/h as the target speed for this experiment, as it is common, and is also the average speed between the speed limits of ordinary roads (60 km/h) and expressways (100 km/h), according to Japan Automobile Foundation [6]. I want to situate the participants in a driving scenario where this speed limit is observed, with an

---

[6] https://web.archive.org/web/20211023150337/https://english.jaf.or.jp/driving-in-japan/traffic-rules, accessed 1 March 2022

additional rule that they need to keep it as close to a specific speed as they could, so that participants would rely on their own speed perception, and I can evaluate participants on the same speed threshold.

As much as possible, I would like to avoid the use of any of the typical indications of speed in cars, like a speedometer reading or its digital numerical counterpart. Apart from additional cognitive load, these indicators rely on current status and physical properties of the vehicle, and not on the perception of the driver.

To restrict the experiment task to just involving the manipulation of driving speed, I created a driving scenario where participants (1) do not have to move the steering wheel to turn left or right, but (2) are required to use the brake and accelerator pedals. To achieve this, I designed the highway as a straight road (to satisfy (1)) but with uneven elevations (to satisfy (2)). If the road is sloping and up and down, participants will be forced to accelerate or decelerate accordingly to still keep the target speed, unlike the scenario where they will just keep their foot on the pedal stationary to keep the same speed on a straight road.

3.3.2 Hypotheses

When I designed this experiment, I formulated the following hypotheses around the absence/presence of the patterns, and the increasing level of information density (which I will interchange with the term ”stimulus density” according to the context). I assume that this would impact not only perceived speed of participants, but also participant’s preferences to either ignore or follow such patterns.

$\mathbf{H_1}$: At least one level of information density has an effect on speed perception, compared to the baseline condition (i.e., no pattern).

For H_1, the augmentation serves as additional stimuli. I assume that with dark environments, participants will not have enough cues to determine if they are maintaining the speed of 80 km/h. While I have removed the different visual cues that participants relied on during initial pilot studies, I assume that participants may still view the trees and background environment as speed or vection indicators.

In my statement of this hypothesis, I would like to emphasize that the existence of effect on at least one pattern is important. On the contrary, if none of the stimulus density levels would have an effect, and their speeds have no statistically significant difference against D_0, then the presence of these augmentations would not have any contribution on influencing speed perception. Thus, I want to state that even if other stimulus density levels may not have an effect, I assume that at least one could potentially have.

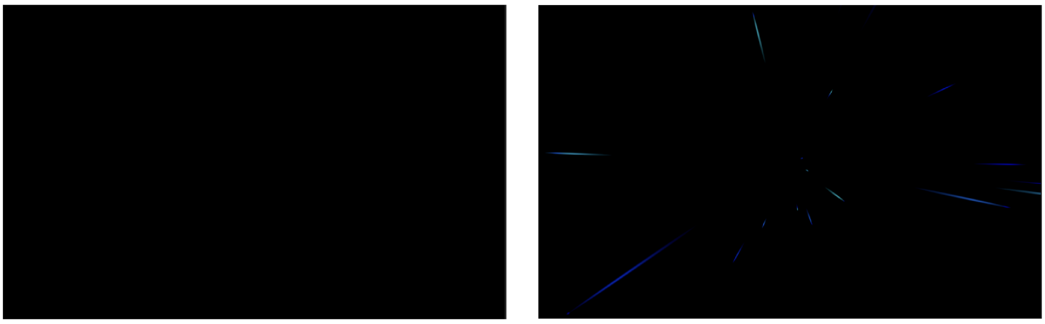
H_2: The pattern with medium density (D_{64}) will give enough speed perception cues without causing too much distraction.

For H_2, I wanted to design the levels of information density and observe the limits of populating the OFOV with AR content as stimuli. I assume that even when presented with a lot of vection cues, participants may find such cues distracting, as it overloads the view with virtual objects. However, I also assume that having more visual cues would still give a higher level of notification for participants, making the stimuli informative. Thus, I hypothesize that the medium level of stimulus density would be a good trade-off. The number (64) and classification (medium) may be arbitrary and ambiguous, however, this may be useful for improving the stimuli in future iterations. The next subsection describes how I selected the levels of stimulus density.

3.3.3 Experiment Variables

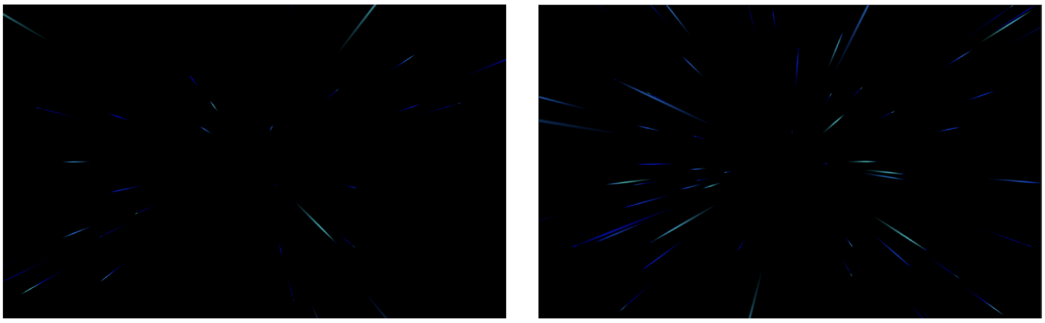
The independent variable of this experiment is the stimulus density (shortened in the Results Section as *Density*) of particles augmented against the real world. The variable is denoted by D_n, where n is the number of particles present on the OFOV with three levels (D_{32}, D_{64}, and D_{128}) and a baseline condition where I show nothing (D_0), even when participants were wearing the OST-HMD. To clarify further, this may seem to be a case of stimulus numerosity or quantity, however, since all the stimuli are moving at a constant rate and within the same area, I have also adopted stimulus density to support the hypotheses statements. Fig. 3.4 show the differences among the three levels.

As I have designed a straight road that varies in elevation in certain segments.



0: No Density Particles (Empty)

32: Low Density Particles (Sparse)



64: Medium Density Particles

128: High Density Particles (Crowded)

Figure 3.4: Screenshots of the three levels of information *Density*. Top row: D_0, D_{32}, Bottom row: D_{64}, D_{128}

I started the road design on flat ground for 2048 m. After that, the road repeats a sequence of an incline-peak-decline-ground. Each upward slope with 2% elevation (incline) continues to a 512 m plateau (peak) at 10 m altitude. Then, the road slopes downward also at 2% before going back to the ground. After another 512 m, the pattern restarts with an upward slope. This is done twice until participants shall see a row of trees blocking the road, signaling the end of each driving task.

I also label the various segments– starting flat road, inclines, peaks, declines, and ground– levels as S_n, A_n, P_n, B_n, and G_n, respectively. The layout of the road according to elevation is also shown in Fig. 3.5.

As I envision this project to actually use an OST-HMD to augment the speed perception-influencing stimuli, I used the Microsoft HoloLens 2.

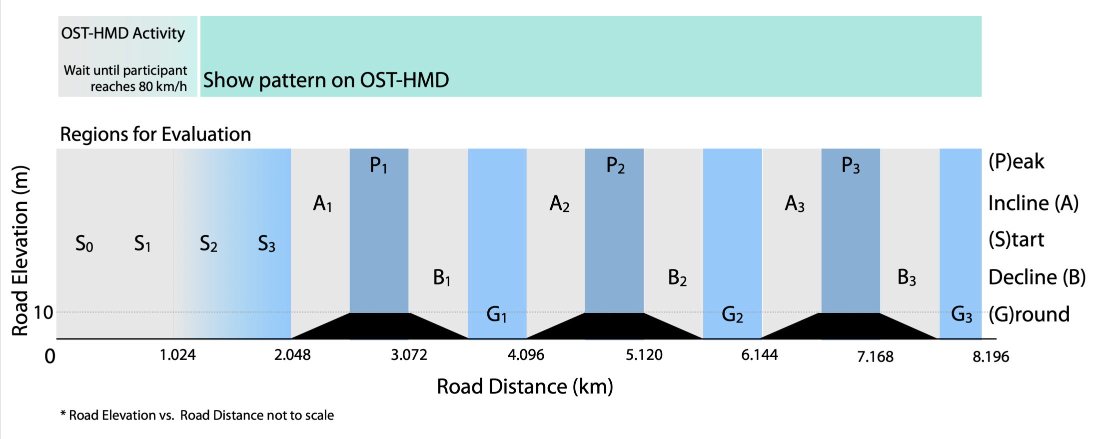


Figure 3.5: Road layout, segmented according to *Elevation*.

3.3.4 Driving Environments

The speed perception-based AR concept would ideally be deployed in the real world environment with a real driving car. However, to ensure safety before the deployment of such an application, this application was developed with driving simulators in a controlled environment.

I designed a highway driving scenario for this experiment. The entire scenario is affected by poor contrast through making participants drive at nighttime with only headlights as light source. I designed it this way so that I can also observe if the same phenomenon of speed misperception happens, and confirm results of related work. I also believe that when the lighting becomes brighter, this condition will impact the contrast of the generated AR content against the bright environment [19].

3.3.5 Design of AR Content

As these types of environments result in driving at faster speeds or an incorrect perception of speed, I want to investigate the possibility of an inverse perceptual process counteracting this incorrect perception.

As this dissertation explores the availability of information within the HMD, I assume that increasing the information density would counteract the lack of stimuli in the environment. Thus, I wanted to create stimuli that would be a reference once they reach a certain threshold speed.

From a pilot study, I have tested various patterns and settled for an animation with a vection pattern. I designed particles that radiate from the center of the screen towards the periphery of the OFOV window, and to the side of the participant. Using the default `ParticleSystem` of Unity 2021.1.25f1, Tables 3.2 and 3.3 show the default values that I used for the design of this animated vection pattern.

I also considered the placement of the AR content against the immediate and effective FOV of the drivers. I augmented the content along the boundaries of the FOV window of the optical see-through head-mounted displays (OST-HMDs). This makes sure that the AR content is not an obstruction and the driver can still see the road ahead. Even with this, I still asked participants if the patterns hindered their focus on driving.

Table 3.2: The particle properties of the `ParticleSystem` used in creating the AR stimuli for this experiment

Particle Properties	**Values**
Duration	3
Start Delay	0
Start Lifetime	0.75-8
Start Speed	10
Start Size	0.034312 (scaled)
Simulation/Playback Speed	8
Rate over Time	1000
Rate over Distance	0
Max Particles	(0, 32, 64, 128) [IV]

Table 3.3: The shape properties of the `ParticleSystem` used in creating the AR stimuli for this experiment

Shape Properties	**Values**
Shape	Hemisphere
Radius	20
Radius Thickness	0
Arc	360°
Mode	Random

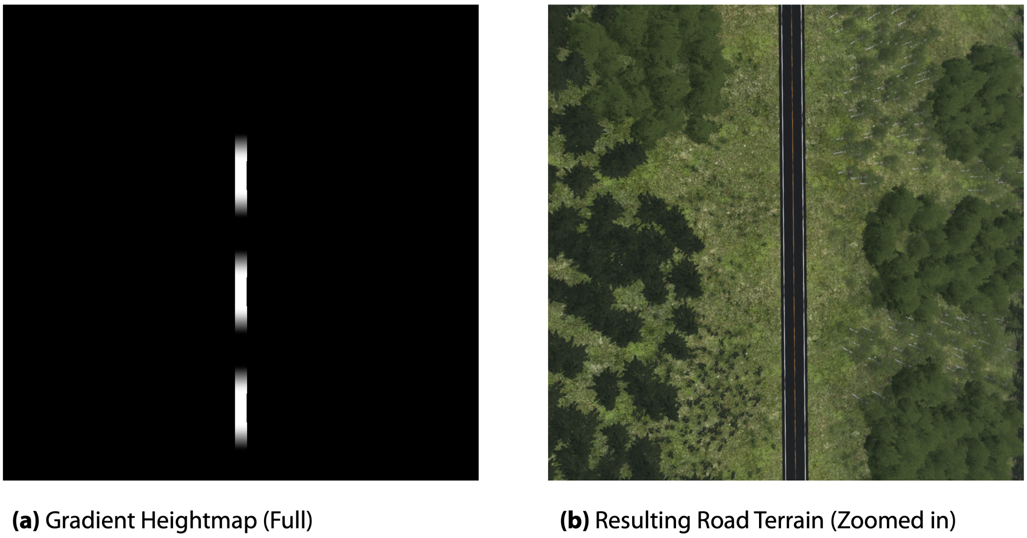
3.3.6 Experimental Setup

The experimental setup in Fig. 3.7 features the use of an OST-HMD in a driving simulator setup. While I envision this experiment to be done in an actual car, I settled first for a driving simulator for safety reasons. In this setup, the participants must be able to view the environment fully within the extents of their horizontal monocular FOV (~200°).

The three monitor setup surrounding a driving seat is similar to the work by Gerber et al. [25]. We also chose 4K resolution for the monitors. This ensures that the driving participant would see the driving scenario with high resolution and wide horizontal FOV. The distance between the surface of the middle monitor and the driving seat where the participant is seated is around 1.3 m. The experimenter sits at the back of the driving simulator setup and observes the participant from another monitor. In this experiment, the monitor

I wanted to verify if viewing augmentations in OST-HMDs would also exhibit this phenomenon of speed perception change. I opted to use this instead of HUDs or windshield projections because near-eye displays entail minimal head and body movement because the view is directly rendered in front of the eyes. In particular, the participants would wear the Microsoft HoloLens 2. I am interested in this setup of monitors and OST-HMDs, since I also want to observe how the lights emitted from the holographic display and from the monitors could still produce a mixed reality simulation. The design of the driving environment should be able to take into account both the monitors' and the OST-HMD's light intensities such that their contrasts do not impact the overall quality of the AR experience.

The participant should be able to adjust the driving speed. The approach is similar to the studies by [65], where the accelerator and the brake pedal would increase or decrease the speed of the vehicle, respectively. While the aim of the experiment is to minimize many interfering stimuli like noise from the driving simulator application or tactile feedback from the driving seat, the steering wheel will still exhibit force feedback for a more realistic driving experience and serve as a notification for any collisions or turbulence in the simulator.



(a) Gradient Heightmap (Full)

(b) Resulting Road Terrain (Zoomed in)

Figure 3.6: The bird's eye view of the terrain maps, as gradient heightmap and the actual rendering used in the experiment.

3.3.7 Design of the Driving Environment

In this experiment I opted to eliminate as many visual cues as possible when designing the scenario. Given the contrast conflict between the monitors and the OST-HMD, I selected a pitch black nighttime driving scenario with monitors' brightness settings set to a minimum. In this case, not only do I address the limitation of the added AR light conflicting with the environment light [19], but I can also present a driving scenario where many important visual cues in the environment are diminished. I asked participants before each start of the experiment proper if they can see the patterns animate and radiate from the center of their view, to which they all said yes.

I used BeamNG.Drive v.0.24[7], a driving game application that also serves as a high quality physics engine for simulated vehicle driving. Additionally, the application also features high customizability through a terrain map editor and an attachable data logging mechanism that details engine activity as well as other important vehicle and driving information.

I used a terrain heightmap that assigns elevation values on a grayscale gradient

---

[7] https://www.beamng.com/game/, accessed 1 February 2022

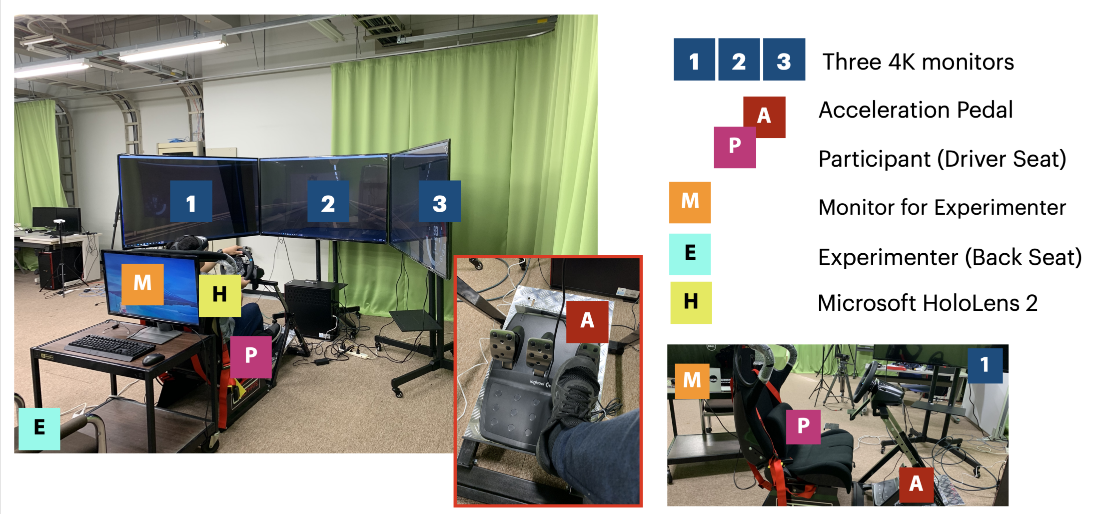


Figure 3.7: The current physical setup consists of monitor displays (1-3), viewed from a driving seat by the participant (P), who controls the speed through pedals for accelerating and braking (A). The experimenter (E) controls the setup behind the participant through a smaller monitor (M).

(Fig. 3.6a). After importing this heightmap, I designed a straight road such that it follows the slopes created by the gradients on the heightmap. From a pilot user study, I have observed that some people follow the rhythm of the broken longitudinal markings found in the middle of the roads, and use them as visual cues for speed perception. So, I had to replace it with a solid longitudinal line in yellow. Finally, I positioned 3D models of forests, trees, bushes, and grass randomly on the roadside.

During the experiment, I also muted the sound from the application, as the pilot user study revealed that the revving sound effects were also being used as speed indicators.

3.3.8 Experiment Procedure

Participants filled out a pre-experiment survey (an excerpt can be found in Fig. B.1) to reveal their experiences with driving and augmented reality, as well as basic demographic information and use of visual implements. For this experiment, I recruited another set of 24 participants from the university through an announcement (10 female and 14 male, 24 to 40 years old (29.13 ± 4.13)).

According to the survey, 21 participants had driving experience, 18 participants had a license to drive. Furthermore, 20 participants had experience using HMDs (including those used in VR). However, only 14 out of all participants had prior experience with AR.

Participants also reported if they had any problems with their vision. Only 5 people reported that they had no problems. Some of them reported having myopia/near-sightedness (12), myopia with astigmatism (3), myopia with photophobia (1), hyperopia/far-sightedness (1), and solely astigmatism (1). Six out of the 24 participants also answered that they have encountered problems relating to dizziness or cybersickness.

However, their physical conditions did not hinder their experiment performance, as everyone successfully finished their experiments until the end without withdrawing. After each trial, I also asked participants if they had any difficulty in viewing the driving scenario, to which they reported no such difficulty. They also wore their visual implements (e.g., eyeglasses, contacts) along with the OST-HMD without any discomfort.

To compensate for varying eye heights of participants when seated, the calibration procedure (Fig. 3.8) requires participants to manipulate the position of a virtual sphere presented on the OST-HMD. Using a keyboard, participants translated the sphere horizontally and vertically such that the center of the sphere and the vanishing point of the road in the driving simulator are aligned.

To further verify the positional calibration, the participant should be able to view what the OST-HMD renders: some particle stimuli that radiates from the same center of the sphere and moves to the edge of the OFOV window. Afterwards, I remove this visualization.

To familiarize participants (especially beginner drivers) with the controls of the application and to train them in perceiving what 80 km/h driving speed is, I started the experiment proper with a tutorial, without wearing any OST-HMD. I also provided them with a digital speedometer that I displayed as a widget in front of the monitor so that they can practice reaching the exact target speed. However, in the experiment proper, I will hide it behind a bookend positioned on the edge of the left monitor, so that only I as experimenter can see the speedometer reading.

The only deliberate instruction for the participants was to keep their speed of

Figure 3.8: The view of the front monitor, featuring the calibration procedure and the blocking of the car speedometer arrow reading from the participant's view.

80 km/h. I give a verbal go signal that they can start driving as soon as I start recording vehicle data and replay footage.

Then, to give the participants an idea about reaching this target speed, I also informed them verbally if they have sustained 80 km/h speed for at least a second, reading it from the experimenter's seat. At this same time, I press a keyboard button to present the pattern for that trial.

To counterbalance order effects, each participant had a unique sequence of all the four conditions (baseline, low, medium, high). Once the participants and I saw the row of trees blocking the road, I also asked them to prepare braking and stopping the vehicle. As soon as the vehicle stops, I also stopped the data recording, and turned off the pattern rendered on the OST-HMD, if there are any.

At the end of each task excluding the baseline, I asked about their level of discomfort, and then asked them to rate their level of confidence in keeping the 80 km/h speed.

Finally, after finishing all four driving tasks, I asked participants to remove the OST-HMD and go out of the driving simulator setup. Then, I asked them several questions for a post-experiment survey, and for feedback regarding the overall driving experience.

The experiment was approved by the institutional review board of the university and took approximately 60 minutes. Participants received a remuneration of JPY 1000 (USD 9) for their time and effort.

Figure 3.9: Average wheelspeed data versus distance from the driving starting point, according to *Density*. At the origin (0 m), all speeds are at 0 km/h.

3.4 Results

The data log file contained information that included the wheel speed, time, and the position of the vehicle along the map. I converted wheel speed data from meters per second (m/s) to kilometers per hour (km/h), and the y-component range of the vehicle's position from [4096, -4096] to [0, 8192] to easily visualize the resulting graphs.

3.4.1 General Differences

Levene's test did not detect a significant difference between the levels of variance across the road segments ($F(3,1532)=2.759$, $p \approx 0.041$ from the start to the end). However, a Shapiro-Wilk test showed that the results were not normally distributed $p<0.001$. Thus, I carried out non-parametric tests for all these observations.

I used Friedman's test to see whether or not there was statistically significant difference among the pattern densities, in general.

Considering the speeds, I did not find any statistically significant difference among the *Density* levels ($\chi^2(3) = 0.85$, $p \approx 0.837$), however, there was a statistically significant difference for *Elevation* from the results of Friedman test, $\chi^2(15) = 164.581$, $p < 0.001$.

While there is no difference among patterns in general, the first hypothesis regarding the influence of the patterns rely on the comparison between the baseline (i.e., no augmentation) and any level of density (i.e., low (D_{32}), medium (D_{64}), high (D_{128})). With this logic, I conducted pairwise Wilcoxon signed-ranks tests with a Benjamin-Hochberg correction for adjusting p values. The results showed no statistically significant difference between D_0 and D_{128} ($p \approx 0.296$). However, the speeds in D_{64} were higher than in D_0 ($p \approx 0.039$). Furthermore, the speeds in D_{32} were higher than in D_0 ($p \approx 0.016$). Thus, the results support the hypothesis H_1, as at least one level of information density has a statistically significant difference from the baseline condition. Fig. 3.10 isolates the average speeds of one level of *Density* against the baseline condition.

However, the same pairwise Wilcoxon signed-ranks tests did not reveal any statistically significant differences for each pair of information densities excluding the baseline condition (D_{32} vs. D_{64}, $p \approx 0.734$; D_{64} vs. D_{128}, $p \approx 0.296$; D_{32} vs. D_{128}, $p \approx 0.286$).

3.4.2 Individual Differences

I also graphed the task performance per individual, which has results that differ to the main graph in Fig. 3.9. These sample individual results, as shown in Fig. B.2, B.3, B.4, B.5 show results of different *Density* levels grouping and coinciding on the same line and different baseline condition measurements.

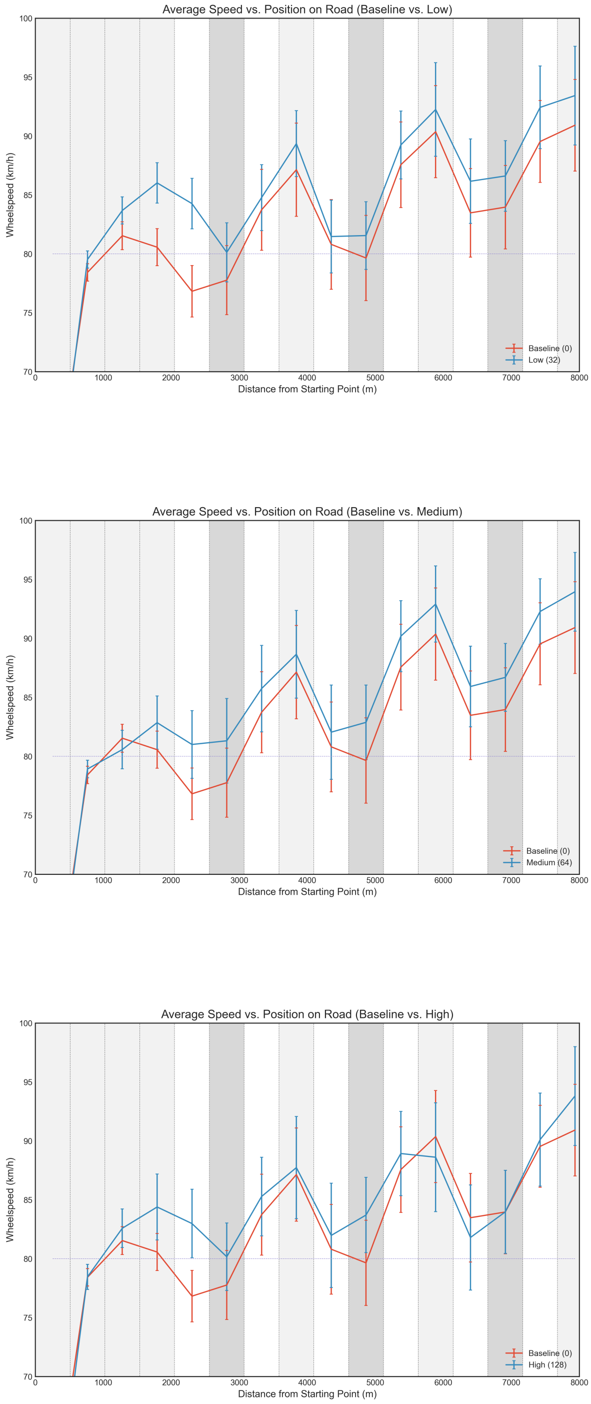


Figure 3.10: Comparing the baseline condition with different levels of *Density* (D_{32}, D_{64}, D_{128}). Error bars represent standard error.

3.4.3 Subjective Evaluations and Preference

After each trial except the baseline condition, participants answered a question about their level of discomfort with the driving scenario, and their level of confidence that they have kept the 80 km/h speed. Table 3.4 shows the average ratings for each level of *Density*.

Participants reported the least discomfort with the sparse density level D_{32}, but rated D_{64} as the condition showing the most discomfort. Furthermore, both *Density* levels were tied at speed perception confidence rating by participants. While participants rated D_{128} the highest in terms of speed perception confidence, all the average confidence ratings were still below 5.0.

At the end of the experiment, I also instructed participants to force rank the three *Density* levels according to their personal preference, without telling them about the differences or providing further criteria for judgment. Participants preferred D_{32} the best (9 participants), and D_{128} the worst (10 participants).

When I asked participants about their preference on whether or not there should be AR stimuli in this driving scenario, they had different answers and reasons. Those who provided reasons (with number of participants indicated in parentheses) looked at various properties of the stimuli.

Half of the participants said in the post-experiment interview that they would prefer not to have any stimuli displayed. Some participants explained that these AR stimuli were distracting (3), unnatural (2), or even sleep-inducing (1). Out of these 12, only one participant clarified that there can be a benefit for over speeding drivers, as it may be a basis for current speed.

Table 3.4: Likert ratings of participants: driving discomfort and speed perception confidence

Factor	D_{32}	D_{64}	D_{128}
Level of Discomfort (1: No discomfort at all; 5: High level of discomfort)	**1.917**	2.250	2.04
Driving Confidence (1: Not confident; 10: Very confident)	4.583	4.583	**4.625**

However, eight participants said that they prefer to have the patterns shown on the display. Some participants qualified which stimuli they preferred (e.g., one preferred D_{128} and possibly D_{64}, another participant liked them both). Other participants expressed how it ”makes the scenario a little lighter [brighter]” (1), “adds sense of speed” (1), and “makes [them] feel that [their] speed is changing” (1). I decided to include the participant who had a tolerant view on the stimuli to this group, saying that “it’s alright to have [the stimuli]”.

Finally, three participants did not directly express any preference. One participant said that there’s “no difference”, and visual cues had no significant effect. One did not notice anything, but still clarified that there were still stimuli presented. And finally, one participant said that they were more worried about keeping the speed, rather than having a preference.

3.5 Discussion

According to the general results of the experiment, I was able to observe a statistically significant difference between the baseline condition and at least one specific level of *Density*. When comparing the three levels of *Density* against each other, I did not observe any statistically significant difference.

In this section, I further analyze the data in terms of both general and individual differences. I also look into insights that may arise when I consider the subjective feedback that participants have given after the end of each task and at the end of the experiment.

3.5.1 Speed Differences According to Stimulus Density

Given that D_{32} and D_{64} show a statistically significant difference against D_0, the results support H_1. I still attribute this to my earlier assumption that the addition of stimuli would serve as additional visual cues for a driving environment that lacks them. However, I would like to qualify that even if the influence is present, it is not the intended influence that I would have expected. For these stimuli to be fully beneficial, they must assist as visual cues that will keep the driving speed on a certain threshold or lower, which in this case was 80 km/h.

For a significant length of the road, the baseline condition (D_0) had the slowest average speed. However, one can observe the exception of the high density condition (D_{128}) having the slowest average speed on the second iteration of the ground to the last incline (G_2, A_3). The graphs that isolate one *Density* level and the baseline condition (Fig. 3.10) show clearer comparisons. These graphs demonstrate how both sparse and medium *Density* levels have higher actual speed values than those of the baseline condition.

The results also imply that the sparser information densities make the participants drive faster. I argue that because the vection was weaker with sparse densities, the discontinuity in optic flow may have made participants consider the speed of these isolated particles. On the other hand, this is not the case for higher densities, where the flow of the patterns is not interrupted, similar to the motion happening on the driving scenario background (with the consistent appearance of trees on the side, and the continued appearance of the road on the center). I would assume that a crowded pattern like D_{128} shows an illusion of uniformity, thus making individual particle speeds less apparent, thus less noticeable. Further experiments with more *Density* levels are required to validate these observations.

The analysis of the main results failed to support H_2. The baseline condition D_0 still had the slowest average speeds, with average values well below the 80 km/h threshold. Thus, it was easy to observe its statistically significant difference against the D_{32} and D_{64} *Density* levels. On the other hand, all the average speed values of D_{32}, D_{64}, and D_{128} even before the initial incline (from S_3 onwards) have already been more than the 80 km/h mark. However, since the comparisons among the average speed values of D_{32}, D_{64}, and D_{128} show no statistically significant difference, the analysis may have been too conservative, or *Density* values outside of this range may show the difference.

The statistically significant difference in speeds among the *Elevation* values also imply that participants were not able to keep a consistent speed. As expected, participants may not be able to accelerate or decelerate back to the target speed once the elevation becomes uneven. The passing of time may also be considered as a reason for this inconsistency, as visual and somatic fatigue may have increased. Maintaining consistent pedal pressure then becomes more difficult.

According to the main results, the higher confidence rating on the high density condition (D_{128}) may also be related to the lack of statistical significance between D_{128} and the baseline condition D_0. However, this relationship between confidence of speed perception and the actual speed results may be weak, as the rating was still low and very close to the two other conditions. Furthermore, additional correlation analysis is needed for this observation to hold.

3.5.2 Individual Speeds and Preferences

I was expecting additional stimuli to regulate the driving speed and keep it at the target speed specified. However, individual results revealed different outcomes every time. The baseline condition D_0, even when it presented no stimulus, still had the worst performance from a few participants (e.g., Fig. B.2). In another individual observation, the opposite is true: a participant's baseline condition had the best estimation of the 80 km/h target speed (e.g., Fig. B.2 and B.5), while other conditions become clear outliers with a speed trajectory that is not consistent with others (e.g., Fig. D_{64} in B.4 and D_{128} in B.5). These various observations that seem to contradict one another shows how each participant had different perceptions of speed even with statistically significant differences in the general results. It would be interesting future work to verify if participants would attribute the same speed towards a specific augmented pattern.

3.5.3 Other Stimulus Properties

In this experiment, I only considered information density as a variable, but with many factors affecting optic flow, there are many ways that this exploration can be extended. For example, the length and lifetime of the particles actually go hand in hand with the density of the patterns on the OST-HMD. If the particles are too short or short-lived, they might not even occupy as much space within the viewing area compared to their opposites that persist on-screen for a longer time. As I have discussed, this may create illusions of discontinued or uniform speed of the optic flow.

How participants also view the patterns as stimuli also matter. Kim and Gabbard [35] explored the distraction potential of AR HUDs, assessing whether the display conveys informative or distracting stimuli for vehicle drivers who may want to use it for assistance. In the context of this experiment, the potential of the vection pattern to be either distracting or notifying participants can also be inferred. The trade-off between a comfortable viewing experience at the expense of incorrect perception is apparent for the sparse condition, D_{32}. It was rated the most comfortable condition out of the three levels of *Density*, even when this sparse condition resulted in having the worst over-speeding values overall.

3.5.4 Experiment Limitations

The limitations to this experiment focus on three things: the selection of stimuli properties, the uniqueness of the task conditions, and the preferences of the participants.

Before this experiment was finalized, there have been multiple iterations of this experiment. On pilot user studies, the patterns vary by animation, brightness and contrast against the environment, as well as the semantic/symbolic load.

Just like in the spatial memory experiment, the selection of the condition levels for the stimuli stemmed from the physical limitations of the system. The limit of 128 particles as the maximum quantity may have prevented the OST-HMD from rendering with latency, therefore slowing down the actual animation of the pattern. However, I argue that this is still not the full density or overcrowdedness that would populate the bounds of the OFOV of the OST-HMD.

Secondly, the task instruction does not allow for any threshold values (e.g., minimum and maximum speeds, and margins of error), as I asked participants to keep exactly at 80 km/h. Even when participants did not fully stop in the middle of the driving scenario, they still reported that estimating this exact speed value put pressure on their driving skills.

Furthermore, while the terrain map had an acceptable percent of elevation well within the usual highway standards, the strict repetition of the inclines and declines is very experimental and may not always mirror that of the real world. While I tried with different types of slope percentages, I also wonder if the variation of slopes would actually alter the overspeeding observed in the general average speed graph (Fig. 3.9).

This design of sloping up and down also created slope illusions [37], which trick the eyes into perceiving inclines as declines, and vice-versa. With this phenomenon, participants may have a hard time assessing whether to accelerate or decelerate, especially when the driving scenario also lacks important visual cues due to its setup.

Throughout the task, the driving scenario was only a simulation, even if 18 out of 24 participants reported that their driving was “natural”. For future work, the environment should be changed from a CG simulation of the driving scenario to a real one, towards a more immersive experience and highly contrasting differences in virtuality (CG augmentations) and reality (real world environment). The results analyzed in this experiment accounted for a scenario where both the augmentation and the environment were emitting artificial lights.

The absence of light in the environment may have also affected some of the participant’s level of fatigue while performing the experiment. Seven participants included “sleepiness” in their post-task and post-survey answers. They cited the regularity or sameness of the scene after prolonged exposure to the poorly-lit scenario, and the absence of other sensory feedback. While this is not the case for the majority, it is still an important factor to emphasize, because it implies danger while driving.

Finally, many participants (12) would prefer not to display patterns against a driving scenario. This implies that participants have not yet grasped the benefits of such view alterations, given that there was an observation of patterns influencing their perception of driving speed.

3.6 Conclusions and Future Work

In sum, I presented an experiment that investigates the influence of overlaying augmented patterns in OST-HMDs on driving, emphasizing visual speed perception of the driving scenario. To provide a demonstrative example, this study focused on the density of augmented information presented on the OST-HMD.

Results suggest the existence of influence towards the driving speed of participants. This means that augmentations displayed against a driving scenario provided additional stimuli or visual cues for participants to consider when driving and keeping a certain speed.

However, it is still not the intended influence that I have expected. In the experiment, two of the patterns (D_{32} and D_{64}) made the driving speed of participants significantly faster than when there was no pattern displayed (D_0). In a practical and realistic viewpoint, this implies that augmenting the patterns I used in this experiment would make drivers overspeed. Additionally, regardless of the presence of the patterns, participants still drove faster on average than the 80 km/h requirement after the first round of elevation change. While I was expecting that at least one of the patterns would minimize this overspeeding, the participants may still have been influenced by the slope illusion.

My main recommendation in extending the study is to observe first which pattern designs and behavior closely approximate the target speed. Participants may view a pre-recorded driving scenario with a constant and consistent 80 km/h, and let them choose among different patterns augmented on the OST-HMD which one is best synchronized with the speed of the scenario. With this approach, participants can be classified according to their selections, or the designs of the patterns can be optimized such that all participants would have similar answers.

As the second hypothesis was not supported, it would be interesting to interpolate or extrapolate information density values from the values that I have used in this experiment, and see where the trade-off between distraction and helpful notification is optimized.

Even when Toui et al. [69] had multiple experiments to show extensive comparisons among various speed patterns, they were still exploring the optimal design of the AR content for velocity recognition, which is similar to the pursuit of this study. There are also so many ways to render animated lines with optic flow, so I suggest further iterating the design of the animations, focusing not only on properties of individual particles, but also the overall optic flow of the patterns when together.

The context of the experiment can also be considered. As driving on a highway seems like an obvious choice, there are many scenarios where influencing speed perception could be beneficial. Perception in other modes of transportation can also be investigated, like on bicycles, on trains, or even on flights. Athletes, coaches, and analysts may find benefit in augmented training for sports, considering an improvement in perception of high-speed objects.

I have also justified the elimination of other sensory information or even additional visual cues so that the analysis can focus solely on the effects of augmentations against a dark driving environment. However, interactions of other factors may produce different outcomes, such as combining these CG augmentations with auditory or tactile feedback, or modifying the lighting and contrast conditions of the driving scenario.

Finally, as I have mentioned earlier as a caveat of this chapter, I conducted the experiment with the intent of only observing the perceptual process of speed. The next step is to explore more ways to use the observations about this perceptual process and apply it to discover the inverse perceptual process of speed.

4. Conclusions and Future Work

In this dissertation, I carried out two experiments that highlight two visual factors that are important in HMD-based AR experiences: the OFOV, and the stimulus density of the vection pattern presented on the display. For each visual factor, I paired them with tasks that involve two major cognitive processes: (spatial) memory and (speed) perception, respectively. On one hand, the perceived limitation of small OFOVs did not significantly matter in recall and transfer tasks. However, the amount of available stimuli can still influence the perception of self-motion in a driving task. I want to conclude by reviewing the main components of the dissertation overview, and recommend future work.

4.1 Interdependence of Cognitive Processes

During the development and after the analyses of these two experiments, I observed how memory and perception as cognitive processes interact and go hand in hand. In the spatial memory experiment, participants needed time to encode and visually perceive the stimuli first before storing them to memory, then move on to memorizing other targets. In the speed perception experiment, it is inherent to have a memory or prior knowledge of what 80 km/h speed looks like before they can proceed with regulating their own perceived speed in the actual task.

But in the context of task performance, there is an additional layer of caution needed to design effective task support tools that use HMDs or OST-HMDs. Regardless of the placement of content within or beyond the OFOV, designers must still take many factors into consideration.

For the spatial memory experiment, I emphasized the importance of the complexity of the content to be memorized. While I have simplified the pair of targets to only differ in terms of locations and colors, there are stimuli in other tasks that involve meanings and order.

As I have mentioned in the setup of the speed perception experiment, I have situated the driving scenario in a nighttime environment to reduce its impact to the contrast of virtual imagery displayed to the user of an OST-HMD. It would be interesting to try out various contrast levels of the driving environment.

The tasks in which OST-HMDs are usually featured as a solution have more complicated structures, including sequences and sub-tasks. The domains in which these tasks belong are also very broad: from surgery procedures in the medical field, to machine assembly and maintenance in industry, to search and rescue operations in crises and disasters. When evaluating these kinds of complicated tasks, there are many cognitive processes that cooperate and depend on each other, all while integrating the assisted input of the OST-HMDs.

4.2 Informative Nature of AR Content as Stimuli

The limiting properties of OST-HMD-based experiences presented challenges on how the AR information is presented.

Due to the limited OFOV, applications that use OST-HMDs should take into account the scale of the space where the content needs to be augmented. In case of AR content outside the OFOV window, visualizations like the wedge should help in guidance. Also, the amount of information that should be available for users to view is also important. Important notification and needless interruption must be balanced so that important stimuli are still rendered and included.

The potential of stimuli to be either informative or distracting [35] is not only relevant in the speed perception experiment. The level of confusion and loss of concentration may arise when the stimuli is distracting. In the spatial memory experiment, I evaluated the use of the wedge visualization for out-of-view targets so that I can assess whether or not they contribute to a faster search of stimuli. In the speed perception experiment, I tested the independent variable of stimulus density with this potential in mind. Whenever stimuli will overcrowd the display, then information conveyed may be superfluous and distracting. On the opposite end, the lack or even absence of visual stimuli will eliminate what is supposed to be an informative display.

4.3 Importance of Motivation for the Task

Human motivation definitely plays a role in task performance, a factor not directly connected to the construction of OST-HMD-based experiences. As I cannot control the different motivations of each participant in both experiments, it would be interesting for future experiments to consider the gravity of task objectives and cognitive load, and how participants will have the same amount of motivation.

This challenge is more apparent in the spatial memorization task, where the participant needs to encode and store spatial information for 15 pairs of targets. Participants' motivation ranged from wanting to get a perfect score to just wanting to get select target pairs correctly memorized.

On the other hand, speed perception and judgments are tricky, as the human visual system needs to process and handle multiple frames of the driving scenario to get a grasp of self-motion for a very specific driving speed value. Participants in the speed perception experiment have reported instances of physical fatigue, e.g., sleepiness, leg strain from pedaling pressure, and eye strain from viewing the scenario for extended periods of time. Even when I have given the option to rest in between tasks and some participants have chosen to rest, the motivation may have already been affected by this reported fatigue.

4.4 Improving OST-HMD Capabilities and Acceptability

In a comprehensive review by Itoh et al. [33], there is still a long way to go before the capabilities of the OST-HMDs match or exceed that of the human visual system. They cited properties like FOV, pixel density, and luminance still need to improve. Some of the observations and motivations of this work originate from the criticisms and limitations of OST-HMDs as an AR device.

In the near future, as long as the prominence of OST-HMD research and development is high, it would be interesting to replicate these experiments with more advanced OST-HMDs. The motivation to broaden the range of applications and improve the overall quality of OST-HMDs will not only improve their versatility, but also their acceptability in daily interactions in the future.

Acknowledgements

Tuloy. Continue. つづく. In any language, I always looked at this word as a challenge, a mantra, a question throughout my post-graduate life. Looking back, I am very thankful for everyone who I have met and interacted with, who showed me in many ways and forms what it means to continue. I was able to live and become this word, and I look forward to the rest of my life with this word in mind.

First of all, I would still like to thank the Government of Japan, through the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the screening committee at NAIST for extending my scholarship under the International Priority Graduate Programs.

I would like to thank my professors for their guidance and support, even when the world is facing a very difficult situation. I look up to all of you as examples of who I want to become, moving forward. I hope to make you all proud.

All of this work will not be possible without the trust that Prof. Hirokazu Kato has given to me. Thank you for believing (in) me since I was an intern in 2013. I will never have believed in myself otherwise. I was able to continue my dreams of pursuing an academic life.

I would also like to thank Asst. Prof. Alexander Plopski for being my direct supervisor for many years. You have taught me the necessary tools for this career (whether in planning, in paper writing, in statistics, in presenting, the list goes on). You continued your support, extending time and effort to ensure that my research is conducted with a high standard.

I would also like to thank the professors of Interactive Media Design Laboratory that I have met during my stay as a student, from the past (Assoc. Prof. Christian Sandor, Asst. Prof. Takafumi Taketomi, Asst. Prof. Goshiro Yamamoto) to present (Assoc. Prof. Masayuki Kanbara, Asst. Prof. Yuichiro Fujimoto, and Asst. Prof. Taishi Sawabe), as well as my dissertation panelist Prof. Kiyoshi Kiyokawa. Your comments about my research are invaluable.

Outside of IMD, I would like to thank Prof. Ma. Mercedes Rodrigo of the Ateneo Laboratory for the Learning Sciences (ALLS) for continued advice about graduate school and academic life. I would also like to thank Assoc. Prof. Matjaž Kljun and Asst. Prof. Klen Čopič Pucihar for accommodating me as an intern at the Humans Interacting with Computers at University of Primorska (HICUP) Laboratory in 2019. It still continues to be one of the best summers of my life.

To the members of Interactive Media Design Laboratory, thank you very much for giving me an environment where I can "outgrow my limits" and continue to grow as a researcher and as a person.

Many thanks to everyone from the International Student Affairs Section, Academic Affairs Section, and Graduate School / Division of Information Science Office, for the continued 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 five years. We have the same Filipino word for "continuing" and "allowing someone home".

Finally and most especially, I dedicate this work to my love, to my dear family. Because of you, I continue to be.

silence: common song
of tilled fields, unlit houses–
red tower lights blink

Takayama (2013)

References

[1] Ryo Akiyama, Goshiro Yamamoto, Toshiyuki Amano, Takafumi Taketomi, Alexander Plopski, Yuichiro Fujimoto, Masayuki Kanbara, Christian Sandor, and Hirokazu Kato. Illusory light: Perceptual appearance control using a projection-induced illusion. *Computers & Graphics*, 91:129–140, 2020.

[2] Patricia L. Alfano and George F. Michel. Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills*, 70(1):35–45, 1990.

[3] Kevin Wayne Arthur. *Effects of field of view on performance with head-mounted displays.* PhD thesis, University of North Carolina at Chapel Hill, 2000.

[4] Richard C Atkinson and Richard M Shiffrin. The control of short-term memory. *Scientific American*, 225(2):82–91, 1971.

[5] Bernard J. Baars and Nicole M. Gage. Chapter 2 - a useful framework. In Bernard J. Baars and Nicole M. Gage, editors, *Fundamentals of Cognitive Neuroscience*, pages 25–57. Academic Press, San Diego, 2013.

[6] Alan D. Baddeley and Graham Hitch. Working memory. In Gordon H. Bower, editor, *Psychology of Learning and Motivation*, volume 8, pages 47–89. Academic Press, 1974.

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

[8] Erica M. Barhorst-Cates, Kristina M. Rand, and Sarah H. Creem-Regehr. The Effects of Restricted Peripheral Field-of-View on Spatial Learning while Navigating. *PloS one*, 11(10):e0163785–e0163785, 2016.

[9] Nicko R. Caluya, Alexander Plopski, Jayzon Ty, Christian Sandor, Takafumi Taketomi, and Hirokazu Kato. Transferability of spatial maps: Augmented vs. virtual reality training. In *Proceedings of 25th IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 387–393, 2018.

[10] Zeynep Cipiloglu, Abdullah Bulbul, and Tolga Capin. A framework for enhancing depth perception in computer graphics. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, APGV '10, page 141 – 148, New York, NY, USA, 2010. Association for Computing Machinery.

[11] Andy Cockburn and Bruce McKenzie. Evaluating the effectiveness of spatial memory in 2d and 3d physical and virtual environments. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '02, page 203 – 210, New York, NY, USA, 2002. Association for Computing Machinery.

[12] Resty C. Collado, Nicko R. Caluya, and Marc Ericson C. Santos. Teachers' evaluation of MotionAR: An augmented reality-based motion graphing application. *Journal of Physics Conference Series*, 1286:012051, aug 2019.

[13] Marco Costa, Leonardo Bonetti, Valeria Vignali, Claudio Lantieri, and Andrea Simone. The role of peripheral vision in vertical road sign identification and discrimination. *Ergonomics*, 61(12):1619–1634, 2018.

[14] Javier M. Covelli, Jannick P. Rolland, Michael Proctor, J. Peter Kincaid, and P. A. Hancock. Field of view effects on pilot performance in flight. *The International Journal of Aviation Psychology*, 20(2):197–219, 2010.

[15] N. Cowan. What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169:323–338, 2008.

[16] C. Diaz, M. Walker, D. A. Szafir, and D. Szafir. Designing for depth perceptions in augmented reality. In *2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 111–122, 2017.

[17] Norman R. Ellis. Automatic and effortful processes in memory for spatial location. *Bulletin of the Psychonomic Society*, 29(1):28–30, 1991.

[18] Barrett Ens, David Ahlström, and Pourang Irani. Moving ahead with peephole pointing: Modelling object selection with head-worn display field of view limitations. In *Proceedings of the 4th ACM Symposium on Spatial User Interaction*, pages 107–110, 2016.

[19] Austin Erickson, Kangsoo Kim, Gerd Bruder, and Gregory F. Welch. Exploring the limitations of environment lighting on optical see-through head-mounted displays. In *Symposium on Spatial User Interaction*, SUI '20, New York, NY, USA, 2020. Association for Computing Machinery.

[20] Ajoy S. Fernandes and Steven K. Feiner. Combating VR sickness through subtle dynamic field-of-view modification. In *2016 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 201–210, 2016.

[21] 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 3rd Augmented Human International Conference*, pages 1–8, 2012.

[22] Hind Gacem, Gilles Bailly, James Eagan, and Eric Lecolinet. Impact of Motorized Projection Guidance on Spatial Memory. In *Proceedings of the 4th ACM Symposium on Spatial User Interaction*, pages 51–59, 2016.

[23] BoYu Gao, HyungSeok Kim, Byungmoon Kim, and Jee in Kim. Artificial landmarks to facilitate spatial learning and recalling for curved visual wall layout in virtual reality. In *Proceedings of the IEEE International Conference on Big Data and Smart Computing*, pages 475–482, 2018.

[24] Nirit Gavish, Teresa Gutiérrez, Sabine Webel, Jorge Rodríguez, Matteo Peveri, Uli Bockholt, and Franco Tecchia. Evaluating virtual reality and augmented reality training for industrial maintenance and assembly tasks. *Interactive Learning Environments*, 23(6):778–798, 2015.

[25] Michael A. Gerber, Ronald Schroeter, and Julia Vehns. A video-based automated driving simulator for automotive ui prototyping, ux and behaviour research. In *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, AutomotiveUI '19, page 14 – 23, New York, NY, USA, 2019. Association for Computing Machinery.

[26] Uwe Gruenefeld, Abdallah El Ali, Susanne Boll, and Wilko Heuten. Beyond halo and wedge: Visualizing out-of-view objects on head-mounted virtual

and augmented reality devices. In *Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services*, pages 1–11, 2018.

[27] Sandra G. Hart. NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9):904–908, 2006.

[28] Abraham M. Hashemian, Alexandra Kitson, Thinh Nquyen-Vo, Hrvoje Benko, Wolfgang Stuerzlinger, and Bernhard E. Riecke. Investigating a sparse peripheral display in a head-mounted display for vr locomotion. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 571–572, 2018.

[29] Lynn Hasher and Rose T. Zacks. Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108(3):356–388, 1979.

[30] Shirin E. Hassan, John C. Hicks, Hao Lei, and Kathleen A. Turano. What is the minimum field of view required for efficient navigation? *Vision Research*, 47(16):2115–2123, 2007.

[31] Ian P. Howard and Brian J. Rogers. *Binocular Vision and Stereopsis.* Oxford Psychology. Oxford University Press, 1995.

[32] Toru Ishikawa. Retention of memory for large-scale spaces. *Memory*, 21(7):807–817, 2013.

[33] Yuta Itoh, Tobias Langlotz, Jonathan Sutton, and Alexander Plopski. Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. *ACM Comput. Surv.*, 54(6), Jul 2021.

[34] Yvonne Jansen, Jonas Schjerlund, and Kasper Hornbæk. Effects of locomotion and visual overview on spatial memory when interacting with wall displays. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, CHI ’19, page 1 – 12, New York, NY, USA, 2019. Association for Computing Machinery.

[35] Hyungil Kim and Joseph L. Gabbard. Assessing distraction potential of augmented reality head-up displays for vehicle drivers. *Human Factors*, 0(0):0018720819844845, 0. PMID: 31063399.

[36] Naohiro Kishishita, Kiyoshi Kiyokawa, Jason Orlosky, Tomohiro Mashita, Haruo Takemura, and Ernst Kruijff. Analysing the effects of a wide field of view augmented reality display on search performance in divided attention tasks. In *2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 177–186, 2014.

[37] Akiyoshi Kitaoka. Slope illusion (magnetic hills) in radan. *Art and Its Role in the History: Between Durability and Transient-ISMS*, pages 751–760, 2015.

[38] Warrunthorn Kittiwongsunthorn, Aleksandar Kovac, and Katsuhiko Kushi. Automotive peripheral vision interface. In *Proceedings of the 5th International Association of Societies of Design Research*, IASDR '13, pages 1–12. 5th International Cogress of IASDR 2013, 2013.

[39] Joshua M. Knapp and Jack M. Loomis. Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators and Virtual Environments*, 13(5):572–577, 2004.

[40] Eric Krokos, Catherine Plaisant, and Amitabh Varshney. Virtual memory palaces: immersion aids recall. *Virtual Reality*, 23(1):1–15, 2019.

[41] Ernst Kruijff, J. Edward Swan, and Steven Feiner. Perceptual issues in augmented reality revisited. In *2010 IEEE International Symposium on Mixed and Augmented Reality*, pages 3–12, 2010.

[42] Mikko Kytö, Aleksi Mäkinen, Jukka Häkkinen, and Pirkko Oittinen. Improving relative depth judgments in augmented reality with auxiliary augmentations. *ACM Transactions on Applied Perception*, 10:1–22, 02 2013.

[43] Cha Lee, Scott Bonebrake, Tobias Hollerer, and Doug A. Bowman. A replication study testing the validity of ar simulation in VR for controlled experiments. In *2009 8th IEEE International Symposium on Mixed and Augmented Reality*, pages 203–204, 2009.

[44] Myungho Lee, Gerd Bruder, Tobias Höllerer, and Greg Welch. Effects of unaugmented periphery and vibrotactile feedback on proxemics with virtual humans in ar. *IEEE Transactions on Visualization and Computer Graphics*, 24(4):1525–1534, 2018.

[45] Eric L.G. Legge, Christopher R. Madan, Enoch T. Ng, and Jeremy B. 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.

[46] Svenja Leifert. The influence of grids on spatial and content memory. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems: Extended Abstracts*, pages 941–946, 2011.

[47] Danielle E. Levac, Meghan E. Huber, and Dagmar Sternad. Learning and transfer of complex motor skills in virtual reality: a perspective review. *Journal of NeuroEngineering and Rehabilitation*, 16(1), October 2019.

[48] Björn Lidestam, Lars Eriksson, and Olle Eriksson. Speed perception affected by field of view: Energy-based versus rhythm-based processing. *Transportation Research Part F: Traffic Psychology and Behaviour*, 65:227–241, 2019.

[49] James Jeng-Weei Lin, Henry B. L. Duh, Habib Abi-Rached, Donald E. Parker, and Thomas III A. Furness. Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment. In *Proceeding of the 9th IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 164–171, 2002.

[50] Coleman Merenda, Hyungil Kim, Kyle Tanous, Joseph L. Gabbard, Blake Feichtl, Teruhisa Misu, and Chihiro Suga. Augmented reality interface design approaches for goal-directed and stimulus-driven driving tasks. *IEEE Transactions on Visualization & Computer Graphics*, 24(11):2875–2885, Nov 2018.

[51] Daniel R. Montello, David Waller, Mary Hegarty, and Anthony E Richardson. Spatial memory of real environments, virtual environments, and maps. *Human Spatial Memory: Remembering Where*, pages 251–285, 2004.

[52] Charles M. Oman, Wayne L. Shebilske, Jason T. Richards, Travis C. Tubre, Andrew C. Beall, Alan Natapoff, and C. M. Oman. Three dimensional spatial memory and learning in real and virtual environments. *Spatial Cognition and Computation*, 2(4):355–372, 2002.

[53] Randy Pausch, Dennis Proffitt, and George Williams. Quantifying immersion in virtual reality. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '97, page 13 – 18, USA, 1997. ACM Press/Addison-Wesley Publishing Co.

[54] Paolo Pretto, Jean-Pierre Bresciani, Gregor Rainer, and Heinrich H Bülthoff. Foggy perception slows us down. *eLife*, 1:e00031, Oct 2012.

[55] Paolo Pretto, Maelle Ogier, Heinrich H. Bülthoff, and Jean-Pierre Bresciani. Influence of the size of the field of view on motion perception. *Computers & Graphics*, 33(2):139–146, 2009.

[56] Long Qian, Alexander Plopski, Nassir Navab, and Peter Kazanzides. Restoring the awareness in the occluded visual field for optical see-through head-mounted displays. *IEEE Transactions on Visualization and Computer Graphics*, 24(11):2936–2946, 2018.

[57] Eric D. Ragan. The effects of higher levels of immersion on procedure memorization performance and implications for educational virtual environments. *Presence*, 19(6):527–543, 2010.

[58] Eric D. Ragan, Doug A. Bowman, Regis Kopper, Cheryl Stinson, Siroberto Scerbo, and Ryan P. McMahan. Effects of field of view and visual complexity on virtual reality training effectiveness for a visual scanning task. *IEEE Transactions on Visualization and Computer Graphics*, 21(7):794–807, 2015.

[59] Eric D. Ragan, Regis Kopper, Philip Schuchardt, and Doug A. Bowman. Studying the effects of stereo, head tracking, and field of regard on a small-scale spatial judgment task. *IEEE Transactions on Visualization and Computer Graphics*, 19(5):886–896, 2013.

[60] Donghao Ren, Tibor Goldschwendt, YunSuk Chang, and Tobias Höllerer. Evaluating wide-field-of-view augmented reality with mixed reality simulation. In *Proceedings of the 23rd IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 93–102, 2016.

[61] Patrick Renner and Thies Pfeiffer. Attention guiding techniques using peripheral vision and eye tracking for feedback in augmented-reality-based assistance systems. In *2017 IEEE Symposium on 3D User Interfaces (3DUI)*, pages 186–194, 2017.

[62] Oscar Rosello, Marc Exposito, and Pattie Maes. Nevermind: Using augmented reality for memorization. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, page 215 – 216, 2016.

[63] Taishi Sawabe, Masayuki Kanbara, and Norihiro Hagita. Diminished reality for acceleration—motion sickness reduction with vection for autonomous driving. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pages 297–299. IEEE, 2016.

[64] Ludwig Sidenmark and Hans Gellersen. Eye&Head: Synergetic eye and head movement for gaze pointing and selection. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*, UIST '19, page 1161 – 1174, New York, NY, USA, 2019. Association for Computing Machinery.

[65] Robert J. Snowden, Nicola Stimpson, and Roy A. Ruddle. Speed perception fogs up as visibility drops. *Nature*, 392(6675):450–450, April 1998.

[66] Fabio Solari, Martina Caramenti, Manuela Chessa, Paolo Pretto, Heinrich H. Bülthoff, and Jean-Pierre Bresciani. A biologically-inspired model to predict perceived visual speed as a function of the stimulated portion of the visual field. *Frontiers in Neural Circuits*, 13, 2019.

[67] Larry R. Squire and Stuart M. Zola. Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences*, 93(24):13515–13522, 1996.

[68] Alexander Toet, Marieke van der Hoeven, Mirela Kahrimanović, and Nico J. Delleman. Effects of field of view on human locomotion. In *Head- and Helmet-Mounted Displays XIII: Design and Applications*, volume 6955, pages 1–11, 2008.

[69] Hayato Toui, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta. Virtual pattern for affecting the velocity recognition of driver. *The IEICE transactions on information and systems*, pages 45–55, 2016.

[70] Christina Trepkowski, Tom David Eibich, Jens Maiero, Alexander Marquardt, Ernst Kruijff, and Steven Feiner. The effect of narrow field of view and information density on visual search performance in augmented reality. In *Proceedings of the 26th IEEE Conference on Virtual Reality and 3D User Interfaces*, pages 1–10, 2019.

[71] 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 ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 3843–3855, 2017.

[72] Christopher D. Wickens, William S. Helton, Justin G. Hollands, and Simon Banbury. *Engineering Psychology and Human Performance*. Routledge, Aug 2021.

[73] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*, pages 143–146, 2011.

[74] Robert Xiao and Hrvoje Benko. *Augmenting the Field-of-View of Head-Mounted Displays with Sparse Peripheral Displays*, page 1221 – 1232. Association for Computing Machinery, New York, NY, USA, 2016.

[75] Jang W. Yoon, Robert E. Chen, Esther J. Kim, Oluwaseun O. Akinduro, Panagiotis Kerezoudis, Phillip K. Han, Phong Si, William D. Freeman, Roberto J. Diaz, Ricardo J. Komotar, Stephen M. Pirris, Benjamin L. Brown, Mohamad Bydon, Michael Y. Wang, Robert E. Wharen Jr, and

Alfredo Quinones-Hinojosa. Augmented reality for the surgeon: Systematic review. *The International Journal of Medical Robotics and Computer Assisted Surgery*, 14(4):e1914, 2018. e1914 RCS-17-0081.R4.

[76] Maryam Zahabi and Ashiq Mohammed Abdul Razak. Adaptive virtual reality-based training: a systematic literature review and framework. *Virtual Reality*, 24:725–752, 2020.

Appendix

A. Spatial Memory Experiment

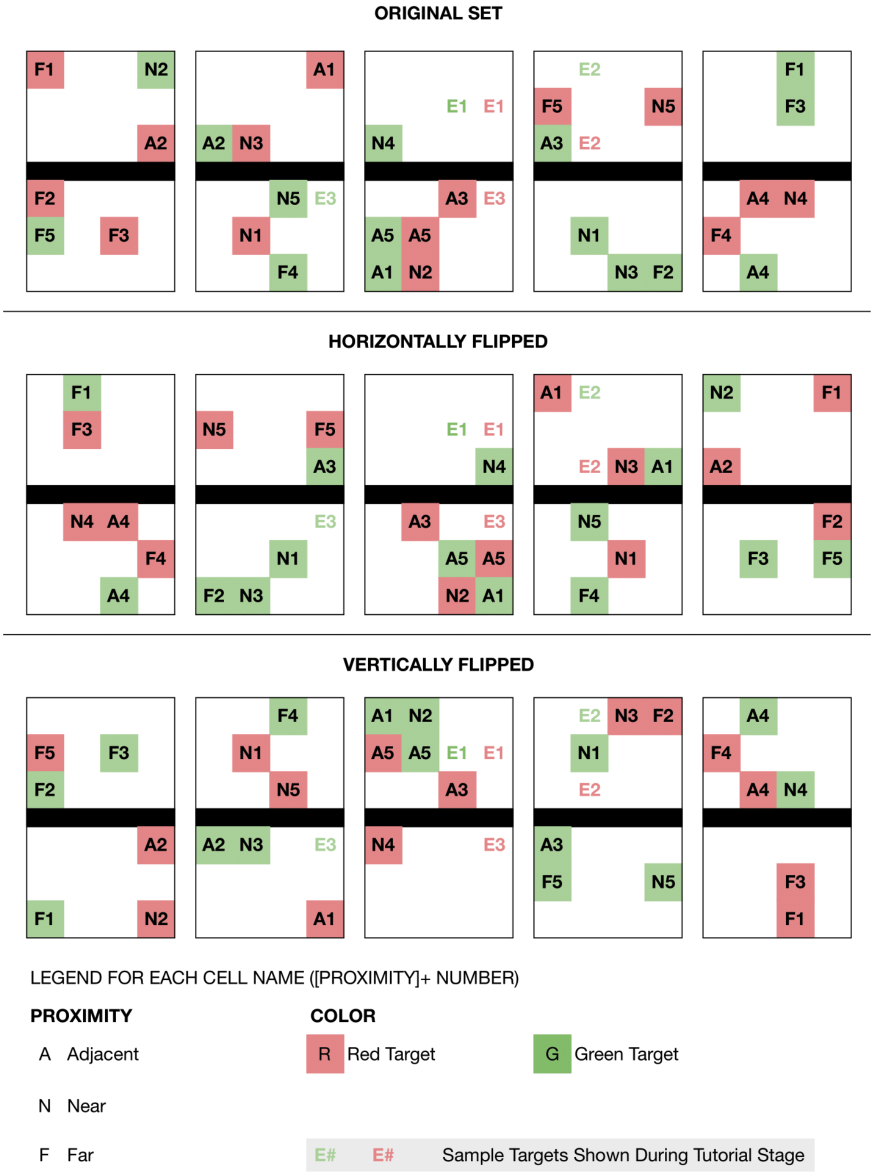
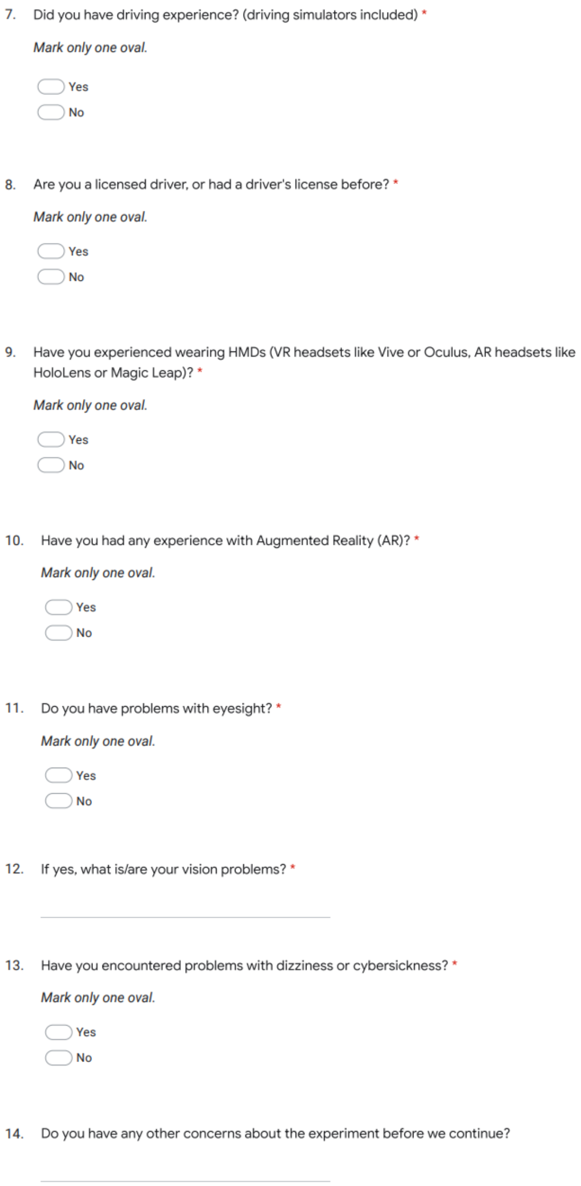


Figure A.1: Diagram illustrating the distribution of target pairs in three *Sets*. Each participant will memorize all *Sets*, and view one *Set* with one *OFOV* level. The colors per set are randomized, except the tutorial stage samples.

B. Speed Perception Experiment



7. Did you have driving experience? (driving simulators included) *

Mark only one oval.

○ Yes
○ No

8. Are you a licensed driver, or had a driver's license before? *

Mark only one oval.

○ Yes
○ No

9. Have you experienced wearing HMDs (VR headsets like Vive or Oculus, AR headsets like HoloLens or Magic Leap)? *

Mark only one oval.

○ Yes
○ No

10. Have you had any experience with Augmented Reality (AR)? *

Mark only one oval.

○ Yes
○ No

11. Do you have problems with eyesight? *

Mark only one oval.

○ Yes
○ No

12. If yes, what is/are your vision problems? *

13. Have you encountered problems with dizziness or cybersickness? *

Mark only one oval.

○ Yes
○ No

14. Do you have any other concerns about the experiment before we continue?

Figure B.1: Excerpt of the pre-experiment survey for the Chapter 3 experiment.

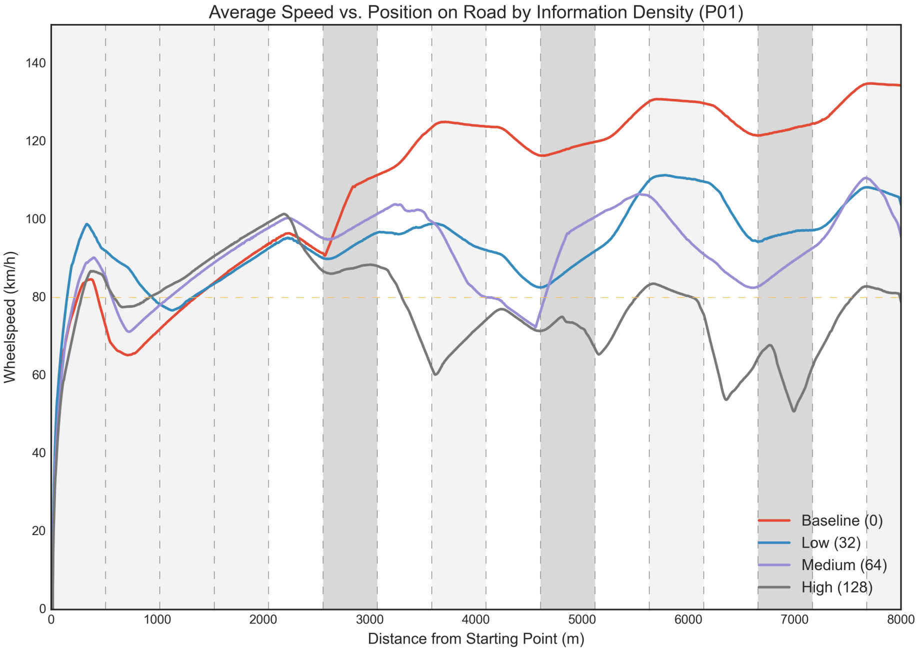


Figure B.2: Sample participant data: significant decrease from overspeeding to target speed.

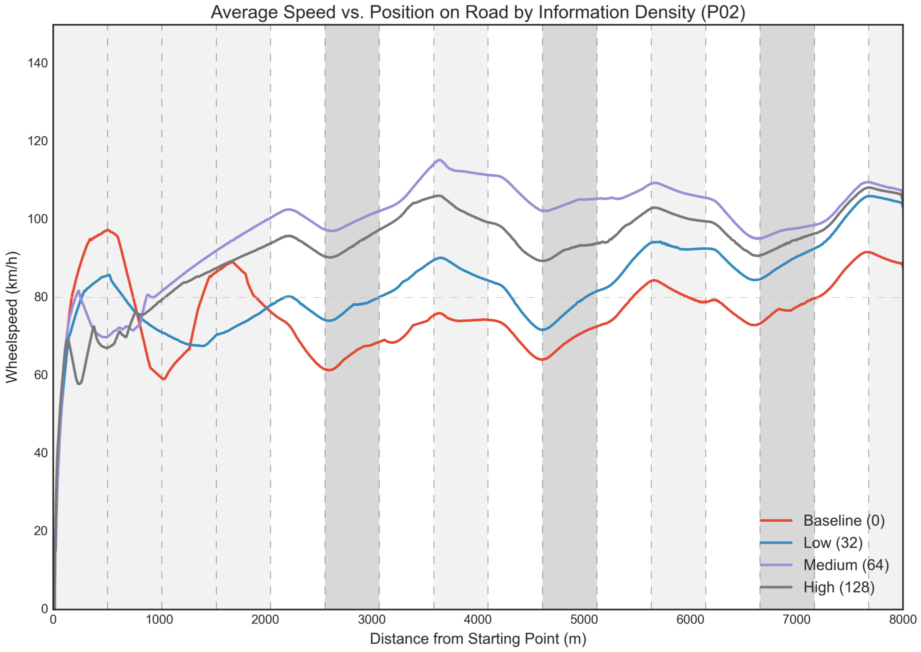


Figure B.3: Sample participant data: patterns influencing overspeeding.

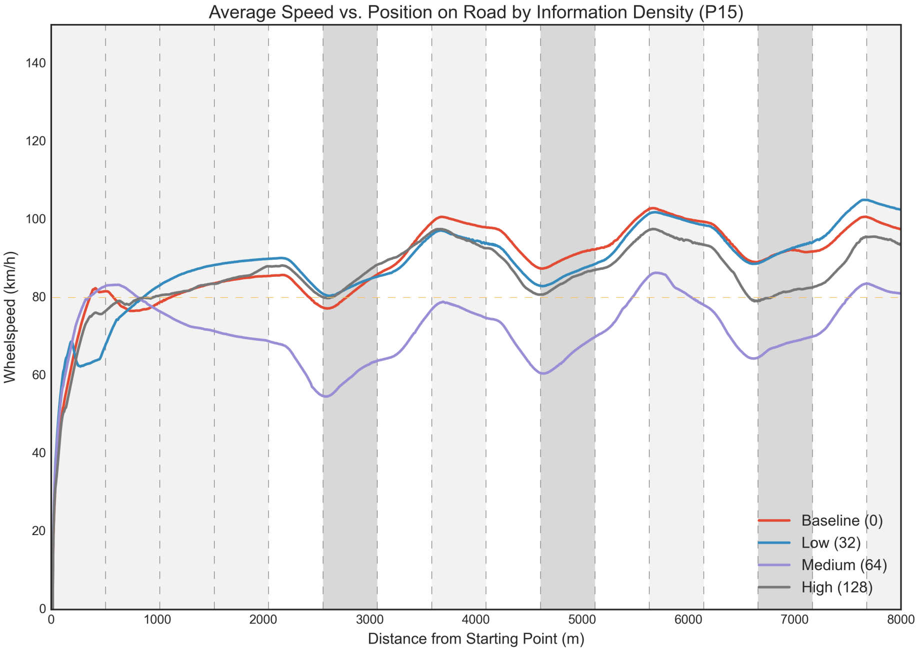


Figure B.4: Medium density (D_{64}) patterns kept driving speed under 80 km/h.

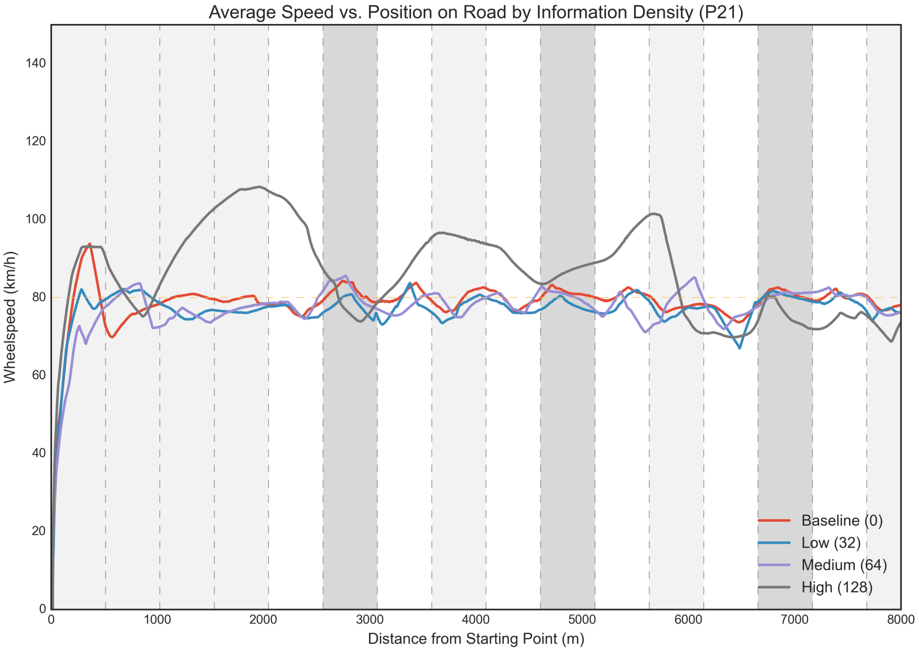


Figure B.5: High density (D_{128}) patterns influence overspeeding more than other densities.

Peer Review Journal Paper

1. **Nicko R. Caluya**, Alexander Plopski, Christian Sandor, Yuichiro Fujimoto, Masayuki Kanbara, Hirokazu Kato, "Does Overlay Field of View in Head-Mounted Displays Affect Spatial Memorization?" , *Computer & Graphics* 102, pp. 554-565, February 2022, Amsterdam: Elsevier. https://doi.org/10.1016/j.cag.2021.09.004. (Chapter 2)

2. Resty C. Collado, **Nicko R. Caluya**, and Marc Ericson C. Santos, "Teachers' Evaluation of MotionAR: An Augmented Reality-Based Motion Graphing Application " *Journal of Physics: Conference Series* 1286 (1), 012051, pp. 1-9, London: IOP Publishing Limited. (Chapter 3)

3. Eric Cesar E. Vidal Jr., Jayzon F. Ty, **Nicko R. Caluya**, and Ma. Mercedes T. Rodrigo,"MAGIS: Mobile Augmented-reality Games for Instructional Support " *Interactive Learning Environments* 27(7), pp. 895-907, 2018, Oxon: Routledge.

4. Maheshya Weerasinghe, Klen Čopič Pucihar, Julie Ducasse, Aaron Quigley, Alice Toniolo, Angela Miguel, **Nicko R. Caluya**, and Matjaž Kljun, "Exploring the future building: Representational effect on projecting oneself into the future office space " *Virtual Reality*, 2022, under review.

Peer Review International Conference

1. **Nicko R. Caluya**, Alexander Plopski, Jayzon F. Ty, Christian Sandor, Takafumi Taketomi. Hirokazu Kato, "Transferability of Spatial Maps: Augmented Versus Virtual Reality Training" , in *Proceedings of the 2018 IEEE Conference on Virtual Reality and 3D User Interfaces* (IEEE VR), 2018, pp. 387-393. (Chapter 2)

2. **Nicko R. Caluya** and Marc Ericson C. Santos, "Kantenbouki VR: A Virtual Reality Authoring Tool for Learning Localized Weather Reporting," in *Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces* (IEEE VR), 2019, pp. 866-867. (Chapter 2)

Other Conference and Workshop Publications or Presentations

1. **Nicko R. Caluya**, Yuichiro Fujimoto, Masayuki Kanbara and Hirokazu Kato, "Influencing Driving Speed Using Perception-Based Augmented Reality", In *Proceedings of the Asia Pacific Symposium on Mixed and Augmented Reality* (APMAR 2021), pp. 1-4. (Chapter 3)