Empirical studies on selection and travel performance of eye-tracking in Virtual Reality

by

(Heather) Yuanyuan Qian

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Computer Science

in

Human Computer Interaction

Carleton University
Ottawa, Ontario

Abstract

We presented two studies on VR selection and travel performances using eye-based interaction via FOVE head-mounted display (HMD). Our selection experiment was modelled after the ISO 9241-9 reciprocal selection task, with targets presented at varying depths in a custom virtual environment. We compared eye-based and head-based in isolation, and the combination of eye-tracking and head-tracking. Results indicate that eye-only offered the worst performance in terms of error rate, selection times, and throughput. Head-only offered significantly better performance. In our travel study, the task involved controlling movement direction while flying through target rings in the air by seven techniques. We found that the completion time and success rates of head+eye were very close to head-only, while eye-only did not perform better than head+eye due to learning effects and calibration issues, which also yield high cybersickness. Head+eye compensated for the eye-tracker issues and would be potentially an alternative to traditional traveling techniques.

Acknowledgements

This thesis would not have been possible without the brilliant guidance, tremendous support and enduring patience of my supervisor Dr. Rob. Teather. I am grateful to him for always guiding me in the right way, using his wisdom to foresee the potential issues in the study, providing me research equipment and pushing me further than I thought I could go. I thank him for teaching me two important courses during the program that enhanced my research capabilities and expanded my field of view. I appreciate his ability to understand my thoughts through my inability to express them explicitly.

I thank Dr. Audrey Girouard for providing useful and professional suggestions to improve the SUI presentation of my selection study. I am grateful to Dr. Sonia Chiasson for embracing me into this program and making all my achievements possible. I thank all the HCI students and colleagues at Carleton University for assisting and participating during my user study sessions.

I am grateful to my husband Thomas for providing me endless help technically and mentally, inspiring me numerous ideas and pushing me against my procrastination. I thank my son Isaac for understanding my work, leaving me much more time for my study and always giving me hugs for encouragement.

I want to thank my parents for being the best parents in the world and for supporting and encouraging me all the time.

Without all of these people, this work would not have been possible.

Table of Contents

List of Tables

List of Illustrations

1 Chapter: Introduction

1.1 Overview

Virtual reality (VR) is a three-dimensional realistic-looking world created by computer graphics for satisfying mankind needs to escape from everyday reality for different reasons and to fulfill human curiosity about exploring beyond the reality [1, 2]. The user in virtual environments (VEs) is not a spectator, but becomes part of this virtual world and is able to interact with the environment. The VR system detects the user's input (gesture, verbal command, etc.) and responses to her in real time. The user receives VR system's feedbacks from sense organs, such as eyes, ears, nose and hands, thus perform a series of actions such as selection, manipulation or navigating in VEs [1, 3]. Like other computer-generated systems, a VR system requires the combination of hardware, software and even sensory synchronicity to provide users immersion and interaction. Recent advances in VR hardware have led to the proliferation of low-cost head-mounted displays (HMDs), such as the Oculus Rift, HTC Vive, and others.

The eye-tracking is the process of measuring either where the eye is focused or the motion of the eye. In the late 19th century, French ophthalmologist Louis Émile Javal observed and measured eye movements during reading [4]. Since then, people have utilized eye-tracker in research on visual systems. In HCI domain, researchers mainly applied eye-tracking in usability evaluation and interactions. Eye-tracker helps researchers investigating usability issues and improving designs (e.g. website) by collecting data like the scanpath and heatspot [5]. Moreover, eye movements can interact with an interface as an input device by dwell, blink, or in combination with other input devices [6– 10]. Many studies have proved that eye-based multimodal interaction would

help target acquisition, selection and manipulation [9– 12]. The use of eye movements is also applicable in virtual reality environments. The most widely used applications is foveated rendering [13, 14], that can enhance the immersion and experience in VR. As many studies [15– 18]investigated the applications of eye-tracking in VR, the significant benefit of eye movements in VR interaction is that it needs much less effort rather than other methods to control point of view or objects to across long distance within large three-dimensional spaces [5]. Therefore, eye-based interaction in VR, which researchers has studied before in various 2D contexts seems a reasonable entry for us to explore.

More recently, the industry has put more interests in research and development of eye-based applications in VR. Tobii, the most famous eye-tracking company, has developed customized eye-tracking solutions for VR HMDs and already integrated with HTC Vive last year[1]. Pupil Labs, a Berlin-based company, that offers open source coding and hackable eye tracking solutions, has launched eye-tracking hardware add-ons for HTC Vive, Microsoft Hololens, and Oculus Rift[2]. Oculus VR, which is the Facebook-backed VR company acquired the Eye Tribe, a firm best known for creating software developer kits that bring gaze-based controls to smartphones, tablets, and PCs[3]. Google acquired Eyefluence, a three-year-old startup that specializes in turning eye movements into virtual actions[4].

---

1 https://www.tobii.com/tech/products/vr/
2 https://pupil-labs.com/vr-ar/
3 https://techcrunch.com/2016/12/28/the-eye-tribe-oculus/
4 https://techcrunch.com/2016/10/24/google-buys-eyefluence-eye-tracking-startup/

11

Illustration 1 Pupil Lab's HTC Vive Binocular Add-on and Microsoft Hololens Add-on

The research object in our research is the FOVE[5]. It incorporates an on-board eye tracker inside of VR HMD that also enables the use of eye tracking as an input technique. Users can interact with the environment via the eyes rather than control a cursor .

Illustration 2 FOVE Head-Mounted Display

Target selection, or target acquisition [19], is a critical user interface task, and involves identifying a specific object from all available objects. As early as 1984, Foley et al. [20] recognized the importance of target selection, and analyzed selection tasks for

---

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

2D GUIs. Since then, many researchers [19, 21, 22, 23] have investigated and evaluated 3D selection in virtual and augmented reality environments. Many innovative selection metaphors emerged, such as the virtual hand [24], ray-casting [21], and image plane interaction [25]. These interaction techniques are based on movement of the hand, or in some cases, the head. Modern virtual reality (VR) systems mostly continue this trend. Head-mounted displays (HMDs) that include a handheld tracked input device, such as the HTC Vive, or Oculus Rift, tend to use virtual hand or ray-based interaction. HMDs that do not include such an input device, such as the Hololens and Samsung Gear VR, instead tend to necessitate the use of gaze direction (i.e., user head orientation) coupled with gestures (e.g., airtap) for interaction. These methods are imprecise, and may yield neck fatigue. Previous 2D selection research has revealed that eye-tracking could offer comparable performance to the mouse, in certain case [24], that eye-tracking may offer compelling alternative to head-based selection.

According to Bowman's classic taxonomy [26], other fundamental VR tasks include manipulation, navigation, system control, and symbolic input. Navigation, in turn, further breaks down into travel (the motor component of moving oneself through a virtual environment) and wayfinding (the cognitive task of route planning through the virtual environment). Travel is a particularly interesting candidate for eye-based interaction. For example, Stellmach and Dachselt [27] conducted a VR travel study but actually conducted a 2D selection task using the eye to select UI elements on a 2D panel thus to control the moving directions. We instead proposed to use the eye as a direct input control for travel via a modified gaze-directed steering.

Gaze-directed steering (travel in the direction the head is looking) is a well-known steering metaphor [26, 28, 29] that has long been used in VR travel since Mine [21] proposed gaze-directed flying approach. Variations on gaze-directed steering are still common today like End Space VR[6] Looking in the direction we move is quite natural; eye tracking offers a more fine-grained approach to this that decouples the movement target from the head-orientation, potentially allowing more natural interaction. The problem with gaze-directed steering is that it couples the view direction and movement vector. However, it also allows users to accomplish an intensive task such as virtual walking or flying at a specified velocity fairly easily. Thus, we were interested in the potential advantages of the combination of head and eye to leverage both benefits.

Above all, our motivations of this thesis originated from the natural mechanism of eye movement and eye-based interactions in various 2D contexts. As many researchers have started to explore the use of eyes in VR and the arousing industrial trends, it is worth to investigate eye-based performance in canonical VR interaction tasks like selection, manipulation and navigation. However, there has been no such studies on the performance of selection and navigation using the eye as a direct control device in VR contexts, our studies would be a good try to see if we could obtain valuable outcomes to inspire the future researches. The main goal of our work was thus to compare the performance of both eye and head-based interaction both in isolation, and in tandem, in both selection and travel tasks. Specifically, our objectives were to answer the following research questions:

---

[6]http://endspacevr.com/

R1: Can the two eye-based techniques (single eye-tracking and the combination of eye-tracking and head movement) perform effectively and efficiently in both selection and travel tasks compared by head-based technique?

R2: How were participants' subjective satisfaction on the two eye-based techniques in both selection and travel tasks?。 。 。

1.2   Contributions

The first primary contribution of our work is the first empirical study based on 3D Fitts' law to evaluate eye- and head-based selection performance in VR. We developed an angular Fitts' law testbed that could benefit other 3D/VR selection studies. The second primary contribution is also the first empirical study of the performance of the eye as a direct control device for travel in VR. Both of these primary contributions compared eye- and head-based performance in isolation from one another, and in tandem, which both revealed the differences between the eye and the traditional head-based performance, as well as explored the possible collaboration of head and eye. A secondary contribution of our work is the comparison of head-based travel and mouse/joystick-based travel, all commonly used travel techniques in VR. We also evaluated the cyber sickness in travel with all input techniques, especially for the eye and head comparison.

1.3   Thesis Outline

This thesis is presented in five chapters. The first chapter introduces the FOVE HMD and our motivations for exploring the eye-based techniques in VR selection and navigation.

The second chapter reviews the prior works related to our research in the area of eye theory and issues. It also covers eye-related interaction in both 2D and 3D/VR, including selection and navigation interactions.

The third chapter provides detail about our eye-based selection study, where we compared selection performance between three eye/head interaction techniques modelled after the ISO 9241-9 reciprocal selection task [30, 31]. This chapter includes hypotheses, descriptions of participants, the hardware, software, experiment design, procedure, results and discussion.

The fourth chapter details our eye-based travel study, where the participants controlled movement direction while flying through target rings in the air by seven input techniques including eye-based techniques. We evaluated the performance and sickness levels. This chapter includes hypotheses, descriptions of participants, the hardware, software, experiment design, procedure, results and discussion.

The final chapter summarizes our findings of this thesis. We discuss the limitations of our research and provide suggestions for future works.

1.4   Associated Publications

Portions of this work have led to the following publications [32] :

1. Qian, Y. Y., & Teather, R. J. (2017, October). The eyes don't have it: an empirical comparison of head-based and eye-based selection in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (pp. 91-98). ACM.

2. Qian, Y., & Teather, R. J. (2017, October). Head vs. eye-based selection in virtual reality. In *Proceedings of the 5th Symposium on Spatial User Interaction* (pp. 151-151). ACM.

2   Chapter: Related work

To implement the test environments, methodologies and measurements, we reviewed two large groups of prior research. The first is eye theory and its applications in 2D and VR, the second part is VR interactions we focused in this thesis.

2.1   Eye Theory and Applications

2.1.1   Eye Theory and Issues

The eyes utilize voluntary and involuntary movement to help acquisition, fixation and tracking visual stimuli. The brain exerts signals through three cranial nerves to control the six extraocular muscles which are attached to the eyeball, thus to control the eye movements [33].

The eyes never stop their movements even when they are fixated at one point. They are always making fast virtually random jittering movements. The photoreceptors and the ganglion cells cannot respond when a constant visual stimulus falls on them. In order to make the image received clearer, the random eye movement keeps changing the stimuli thus makes photoreceptors and the ganglion cells being active [34]. These short and rapid movements that occur when the eyes are scanning an area are referred to as saccades. The eyes move as fast as they can during a saccade with a typical duration of 200 milliseconds (ms), but the speed is not consciously controlled. It is useful to scan an area with the fovea of the eye in a high resolution [35].

The fovea is a small area of the retina which covers a one-degree angle in people. It enables us to see the object more accurate. When we are watching a moving object or pursuing it, the head also moves to assist in tracking. But the process of visual information by head movement cannot catch up a fast movement [36]. In order to see the

moving object clearly, the eyes move as well and try to aim the object image on the fovea. Lanman et al. [37] conducted experiments using trained monkeys, comparing eye and head movements when tracking moving objects. They report that head movement closely followed the target, while the eye gaze vector was relatively close to the head vector, but moved somewhat erratically due to saccades. Despite the irregularity of individual eye and head movements, their combination allowed precise target tracking, regardless if the head position was fixed or free. The authors argued that the vestibular system coordinated eye and head motion during tracking, yielding smooth pursuit. These results support our hypothesis for our first study (Chapter 3) that our selection technique employing both the eye tracker and head motion should perform at least as well as head motion, while eye-based selection should have the worst accuracy. They would also support our hypothesis that our travel technique of head and eye should perform better than the eye-tracking and head motion for our second study (Chapter 4).

2.1.2   Eye-based Interaction and Issues in 2D

Researchers have explored the use of eye-tracking since 1980s [37, 38]. A recurring issue in eye-based interaction is the so-called "Midas Touch" problem: subtle unconscious eye movements can yield unintended consequences, since eye-input is "always on". In 1990, Jacob [6] has investigated eye blink and dwell time as selection mechanisms in an effort to overcome this issue. Since then, researchers found that using a combination method could solve the problem. There have been many interaction techniques employing gaze along with another technique, such as hand-input, hand gestures. For example, Rozado and Mardanbegi [8] utilized gaze to identify objects and then control the objects by hand gesture. Chatterjee et al. [9] extended the categories and

18

approaches of gaze+gesture interaction, and report that this interaction could have a very close performance of mouse or keyboard. In our selection study, we also avoid this issue by requiring users to press a key on the keyboard to indicate selection. Stellmach et al. [10] combined gaze with touch and tilt to manipulate pictures. She reported that gaze-directed pivot zoom in combination with a mouse or a gesture would be a good alternative to zoom or investigate details of information [11]. She also proposed that gaze-supported multimodal interaction would help target acquisition, selection and manipulation [11, 12]. In addition, other studies investigated gaze+touch for remote rotate-scale-translate tasks, pan and zoom, and interaction on tablets [39– 42]. Gaze can even be used with another hand-free input technique, like foot-based input, to interact with desktop computers [44]. Mardanbegi et al. [15] was the first to investigate eye-based interaction working with head-mounted device to infer head gestures. Their results showed that some gestures were reliable thus implied that the collaboration of eye and head should be easily understood by users and might have a better performance. Again, in our experiment, we hypothesized that the combination input of eye and head should not perform worse than the single ones.

2.1.3   Eye-based Application in 3D/VR

Many researchers have noted the possibilities of using eye-tracking in virtual reality. Several studies [15, 16, 17, 44] employed eye tracking for applications other than interaction tasks. In 1990, Starker and Bolt [16] used eye-tracking to monitor and analyze user interest in three-dimensional objects and interface. More recently, Essig et al. [17] implemented a VICON-EyeTracking visualizer, which displayed the 3D eyegaze vector from the eye tracker within the motion-capture system. In a grasping task, their system

performed well for larger objects but less so for smaller objects, since a greater number of eye saccades occurred towards boundaries. In testing a variety of objects, they found that a sphere yielded the best results as assessed in a manual annotation task. This was likely because the sphere was bigger than a cup and stapler object, and was not occluded during grasping. These results are consistent with the selection literature, which outlines the importance of target size, distance [19], and occlusion [46].

We also look to studies using eye tracking in 3D games, which in some ways, are similar to VR. For example, Isokoski and Martin [47] conducted a study using a first-person shooting game controlled by eye-tracking. Their goal was to evaluate the eye-tracker's efficiency in game shooting tasks. They pointed out that eye-based input might be an alternative to the traditional mouse+keyboard. Smith and Graham [48] explored the eye-tracker as a control device in several video games, i.e., a first-person shooting game, a role playing game and an action/arcade game. Notably, they utilized eye-tracking to control view orientation in the FPS game, similar to our eye-only travel technique. They reported that although the eye performed slower than the mouse, the intuitive interactive way of eye-tracking increased immersion and significantly enhanced game experience.

It is worth noting other applications of eye tracking in immersive VR. Ohshima et al. [49] implemented a gaze detection technique in VEs. Duchowski et al [18] applied binocular eye tracking in virtual aircraft inspection training by recording participants' head pose and eye gaze orientation. Steptoe et al. [50] presented a multi-user VR application displayed in a CAVE. They used mobile eye-trackers to control user avatar gaze direction, with the intent of improving communication between users. They report that participants' gaze targeted the interviewer avatar 66.7% of the time when asked a

question. However, eye tracker noise created some confusion as to where participants were looking, contributing to 11.1% of ambiguous cases. We anticipate eye tracker noise may similarly affect our results.

2.1.4   Eye-based Selection in 2D

There have been many studies on eye-related selection in 2D settings. For example, research on eye-only selection conducted by Sibert and Jacob [51] revealed that eye gaze selection was faster than using a mouse. Their algorithm could compensate for quick eye movements, and could potentially be adapted for use in virtual environments. They report that there is also physiological evidence that saccades should be faster than arm movements, which may explain their results. This reinforces our opinion that eye-tracking may prove a useful interaction paradigm in VR. Fono and Vertegaal [52] compared four selection techniques, and report that eye tracking with key activation was faster and more preferred than a mouse and keyboard.

However, the most common and empirical methodology to evaluate selection performance is Fitts' law. Fitts' law is a predictive model that describes the speed-accuracy tradeoff in pointing tasks. It predicts that the time required to rapidly move to a target area is a function of the ratio between the distance to the target and the width of the target [53]. It implies that the further or smaller a target is, the harder it will be to select.
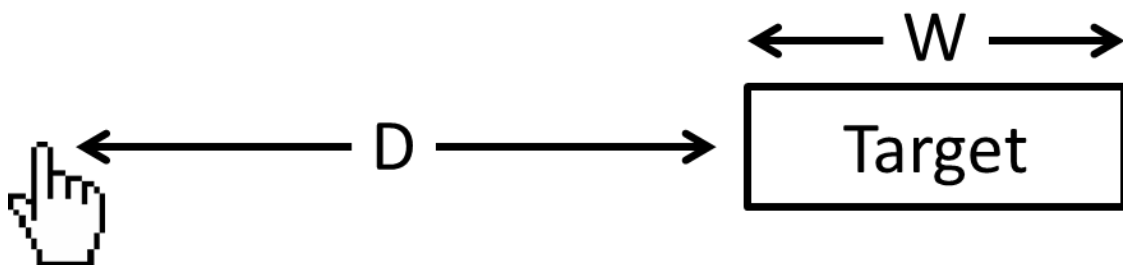
Illustration 3 Fitts's law Demonstration

The model is as below:

$$MT = a + b * ID = a + b * \log_2\left(\frac{D}{W} + 1\right) \qquad \text{Eq.1}$$

where MT is the average time to complete the movement. a and b are constants that depend on the choice of input device and are usually determined empirically by regression analysis. ID is the index of difficulty. D is the distance from the starting point to the center of the target. W is the width of the target measured along the axis of motion. Equation 2 is the recommended method of computing throughput in ISO 9241-9 [54].

$$TP = \frac{IDe}{MT} \qquad \text{Eq.2}$$

where IDe is the effective index of difficulty derived by the effective distance and width. Throughput IP is the index of performance (in bits per second).Fitts originally developed Fitts' law for one-dimensional pointing tasks, but it has been successfully adapted to 2D pointing tasks[54, 55]. Vertegaal conducted a Fitts' law evaluation of eye tracking [57] and found that eye tracking with dwell time performed best among four conditions (mouse, stylus, eye tracking with dwell, and eye tracking with click). However, this study did not employ the standardized methodology for computing throughput (incorporating the so-called accuracy adjustment), the resultant throughput scores cannot be directly compared to other work. Notably, the eye tracker also suffered from a high error rate for both selection methods. MacKenzie presented an overview of several issues in using eye trackers for input with Fitts' law [7]. He also presented the results of two experiments investigating different selection methods using eye tracking, including dwell length, blink, and pressing a key. The eye tracking conditions yielded throughput in the range of 1.16 bits/s to 3.78 bits/s. For reference, ISO-standard compliant studies typically

report mouse throughput of around 4.5 bits/s. In these studies, MacKenzie reported mouse throughput of around 4.7 bits/s.

2.2   Interaction in VR

Interaction techniques are methods that we use to accomplish a given 3D/VR interaction task via the interface and that they include both hardware and software components. Selection, manipulation and travel are the most common and fundamental interaction techniques in the majority of 3D and VR user interfaces. The other two tasks, system control and symbolic input have not been as heavily researched as above three, but they are nonetheless important for many 3D UIs [58].

2.2.1   Selection in VR

Selection techniques include exocentric metaphors and egocentric metaphors. Egocentric metaphors such as virtual hand and ray-based metaphors [24] are in widespread usage today. The virtual hand technique is a metaphor to select an object intersected by hand/tracker a VE, but it is hard to select remote objects. The ray-based metaphor is designed to address the problem. It casts a ray through the cursor or from hand, then selects the first object intersected [58]. Image-plane selection [25] is another "compromise" technique, that supports 2DOF selection of remote objects by touching and manipulating objects' 2D projections on a virtual image-plane located in front of the user. Lee et al. [59] compared image-plane selection to a hand-directed ray, head-directed ray, and a ray controlled by both the head and hand, and report that image-plane selection performed best. Using the eye-tracking capability of a selection technique is perhaps closest to image-plane selection [25]. It requires only 2DOF input, since the user must only fixate on a given pixel. We note that from a technical point of view, this still results

in ray-casting, similar to the use of a mouse for 3D selection. In contrast, head-based

selection uses 6DOF movement of the head –although through careful and deliberate

head movements, a user could constrain this to just 2DOF rotation. We thus anticipate

that eye-tracking could offer superior performance.

Teather and Stuerzlinger [23, 59] extended the ISO 9241-9 standard for

evaluating 3D selection techniques in 3D contexts, using various target depth

combinations, and was validated using both a mouse (for consistency with 2D studies

using the standard) and various 3D tracking devices. Because of the similarity to 2DOF

interaction when using the eye, it was a reasonable methodology to follow the pure

selection task with IOS standard. Hence, an attractive feature of our work was presenting

a plausible 3D task, but supporting comparison between 2D and 3D selection techniques.

We know that the 3DOF movements in 3D/VR requires rotations as well as

translations due to the extra dimension. The early study [61] proposed that rotational

movements have a similar IP as translational movements. After decades of 2D Fitts' law

based studies, several researchers [30, 31] derived modified 3D Fitts' law. They

calculated *ID* using rotation angle between targets for distance, and the angular size of

the target. *ID* was calculated as follows:

$$ID = \log_2 \left(\frac{\alpha}{\omega} + 1 \right)$$
<div align="right">Eq.3</div>

where α is the rotation angle and ω is the angular size the target.

2.2.1.1   Eye-based Selection in 3D/VR

A few researchers have investigated eye-related VR pointing. In 2000, Tanriverdi

and Jacob [62] compared gaze-based pointing and hand-based pointing among

geometrical objects in VR environment. They found the gaze-based pointing was much

faster than hand-based pointing, especially for distant objects. Interestingly, the authors

mentioned that the traditional 2D Fitts' Law was not applied for eye-based selection and

the moving distance should be related to the performance. However, they might somehow

inspired researchers to derive modified 3D Fitts' Law ten years later.

In 2003, Cournia et al. [63] extended Tanriverdi and Jacob's study and found the

contrary result, i.e., gaze-based pointing performed significantly worse than hand-based

pointing. To avoid unfair and different experimental setting, our selection experiment

implemented all selection techniques in a relatively standard 3D Fitts' Law environment.

Thus, this was the first empirical study on eye-based selection in VR.

2.2.2   Navigation in VR

Navigation includes both travel and wayfinding [64]. Travel is the motor

component of navigation—the low-level actions that the user takes to control the position

and orientation of his viewpoint. In the real world, travel is the more physical navigation

task, involving moving feet, turning a steering wheel, letting out a throttle, and so on. In

the virtual world, travel techniques allow the user to translate and/or rotate the viewpoint

and to modify the conditions of movement, such as the velocity. Wayfinding is the

cognitive component of navigation—high-level thinking, planning, and decision-making

related to user movement. It involves spatial understanding and planning tasks, such as

determining the current location within the environment, determining a path from the

current location to a goal location, and building a mental map of the environment [58].

Our second study (chapter 4) focused on a "raw" travel task since travel is more

influenced by input techniques, whereas wayfinding has more to do with environment

design and visual aids. The second study (Chapter 4) compared the performance between single and combination input devices of eye, head, mouse and joystick.

There were a few relevant studies on travel test environments, techniques and evaluations that we used for inspiration here. Nelson et al. [65] conducted a virtual flying study to evaluate their brain-body-actuated control. They had two tasks: the first was to fly through hoops and as close to the center of the hoops as possible, the second was that there were ribbons connected between the hoops, they should fly within the boundaries of the hoops. Their post-test questionnaires were NASA task load index (TLX) and modified simulator sickness questionnaire (SSQ) [66]. We modeled our travel task after this, and we employed similar qualitative measures (e.g., NASA-TLX and SSQ questionnaires).

Cybersickness is similar to the motion sickness symptoms during or after exposure in a virtual environment [67]. The conflicts between three sensory systems: visual, vestibular and proprioceptive would primarily produce cybersickness [67, 68]. Thus the cycbersickness would occur when we are utilizing our visual and vestibular system (eye and head) to travel in VR. Hettinger et al. [70] indicated that a fixed-based visual display produced vection and sickness. When there is a significant mismatch between visual information and vestibular information (as is usually the case in VR travel supported by joysticks), people tend to experience motion sickness. So et al. [71] investigated the cybersickness levels by navigation speed and found that the speed significantly influenced on the levels of vection in the first 5 min and the level of nausea after 10 min from the beginning of the exposure to a VE. The effects became insignificant after 10 and 30 min respectively. In our pilot study, each continuous session

of our experiment lasted less than 6 min, which implied that certain level of vection might occur, but participants should not experience an uncomfortable degree of nausea. They reported that the increasing speed caused the increasing levels of vection and sickness. Learning from this past research and our pilot testing, we picked a fixed speed level to offset cybersickness in the experiment. Based on these results, we expect that head-only would yield lowest levels of cyber sickness. Similarly, our three combination travel techniques should yield lower cyber sickness levels than their corresponding single inputs, due to consistent visual-vestibular information provided by head-tracked viewpoint control.

Moreover, Chen et al. [72] compared head-based and joystick-based navigation techniques. They concluded that the head-based paradigm was superior to the joystick on user performance, presence and cyber sickness. In our travel study (Chapter 4), we thus hypothesized that the joystick would have lower performance and user satisfaction than head-only, and in the combination comparison, the head+joystick should have the worst performance among the three combination input techniques.

2.2.2.1   Eye-based Travel in 3D/VR

Likely the closest study to our travel study (Chapter 4) is that of Stellmach and Dachselt [27]. They investigated eye-based input for steering in virtual environments. The participants needed to get a target position from a start position in 5 different difficulty levels. In order to complete the tasks, participants had to use their eyes to perform rotations and translations by looking at a 2D UI. They found that the continuous gradient-based input, which was one of the eye-based input variables, offered the fastest completion time and was most preferred by participants. Their post-test questionnaire

employed Bowman's traveling questionnaire [26], which we also used in our experiment. In our study, we enabled in-air 6DOF travel (e.g., flying by looking in the intended directly), and looked at how precisely participants could fly through a trail of rings. We also evaluated the cyber sickness caused by eyes other techniques.

3   Chapter: Selection Study

We conducted an experiment based on the international standard, ISO 9241-9 [54], which utilizes Fitts' law [53] to evaluate pointing devices [73]. We compared three different selection techniques using the FOVE: 1) eye-based selection without head-tracking, which we dub eye-only selection, 2) head-based selection without eye-tracking, dubbed head-only selection, and 3) eye-tracking and head-tracking enabled at the same time, henceforth eye+head selection. We compared these selection techniques across several different combinations of target size and depth and assess effectiveness of the eye as a selection method, especially in comparison to head-based selection common to VR systems

3.1   Hypotheses

Our hypotheses included:

H1: Eye+head would offer the best of speed among the three selection techniques, because humans are already well-adapted to coordinating eye and head movement [37].

H2: Head-only would offer the lowest error rate, due to the inherent imprecision of eye-tracking [37].

H3: Participants would prefer eye+head over the other two selection techniques since it leverages the advantages of both head- and eye-only selection.

3.2   Participants

We recruited eighteen participants (aged 18 to 40, μ=28 years, 12 male). All participants were daily computer users (μ=5 hours/day). None had prior experience with eye tracking. Half (nine) had no prior VR experience, five had limited VR experience (having used it once or twice ever), and the rest used VR an average of 5 times in a

month. All participants had colour vision. Fourteen had normal vision, four participants had corrected visions, they wore contact lenses or glasses. All could see stereo, as assessed by pre-test trials. One potential participant did not pass the calibration, then we terminated the pre-test trail.

3.3   Apparatus

Participants wore a FOVE HMD in all trials. See Illustration 4. Participant wearing the FOVE HMD while performing the task. The FOVE display resolution is 2560 x 1440 with a 100° field of view. A unique feature of the display is the two integrated infrared eye-trackers, which offer tracking precision of better than 1° at a 120 Hz sampling rate. Like other HMDs, the FOVE offers IMU-based sensing of head orientation, and optical tracking of head position. However, it does not offer IPD correction.

Illustration 4 FOVE Head-Mounted Display in the experiment

The experiment was conducted on a desktop computer, with an Intel Core i5-4590 CPU, an NVIDIA GeForce GTX 1060 GPU, and 8GB RAM. The experimental interface and testbed was based on discrete-task implementation of the multi-directional tapping test in ISO 9241-9. The software presented a simple virtual environment with spherical

targets displayed at the specified depth. See Illustration 5. The software was developed using Unity 5.5 and C#.

3.4   Procedure

The experiment took approximately 40 minutes in total for each participant. Participants were first briefed on the purpose and objectives of the experiment, then provided informed consent before continuing.
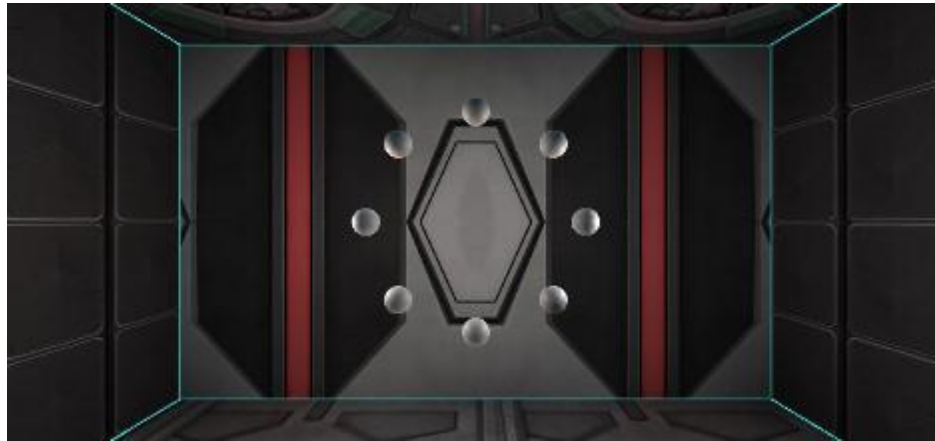
Illustration 5 Software used in the experiment depicting the selection task.

Upon starting the experiment, participants sat approximately 60 cm from the FOVE position tracker, which was mounted on the monitor as seen in Illustration 4. They first completed the FOVE calibration process, which took approximately one minute. Calibration involved gazing at targets that appeared at varying points on the display. This calibration process was also used as pre-screening for the participants: Prospective participants who were unable to complete the calibration process were disqualified from taking part in the experiment. Prior to each new input technique using the eye tracker (i.e., eye-only and eye+head), the eye tracker was re-calibrated to ensure accuracy throughout the experiment. Following calibration, the actual experiment began.

The software presented eight gray spheres in circular arrangement in the screen centre. See Illustration 5. Participants were instructed to select the orange highlighted sphere as quickly and accurately as possible. Selection involved moving the cursor (controlled by either the eye tracker or head orientation) to the orange sphere and pressing the "z" key. The participant's finger was positioned on the "z" key from calibration to the experiment's end to avoid homing/search for the key. Alternative selection indication methods would also influence results (e.g., fixating the eye on a target for a specified timeout would decrease selection speed and thus also influence throughput [74]). We note here that Brown et al. [75] found no significant difference between pressing a key and a "proper" mouse button in selection tasks. However, our future work will focus on alternative selection indication methods.

Upon completing a selection trial, regardless if the target was hit or missed, the next target sphere would highlight orange. A miss was determined by whether the cursor was over the target or not when selection took place. Software logged selection coordinates, whether the target was hit, and selection time. Upon completion of all trials, participants completed a 7-point Likert scale questionnaire based on ISO 9241-9 and were debriefed in a short interview.

3.5   Design

The experiment employed a 3 × 3 × 4 within-subjects design. The independent variables and their levels were as follows:

Input Method: Eye-only, head-only, eye+head

Target Width: 0.25m, 0.5 m, 0.75 m

Target Depth: 5 m, 7 m, 9 m, mixed

With eye-only, the FOVE head tracker was disabled. A ray cast from the eye intersected with the targets directly. With head-only, the FOVE eye tracker was disabled. The orientation of the head derived from the head tracker. The cursor was fixed in the screen centre. A ray cast from head orientation intersected with the targets. The eye+head input method used both the eye and head trackers, and represents the "default" usage of the FOVE. The head orientation controlled the point of view, while a ray cast from the eye intersected with the targets. Although eye-only does not represent typical usage of the FOVE, it was included to provide a reasonable comparison point to previous eye-tracking Fitts' law studies [7].

Three target sizes yielded three distinct indices of difficulty, calculated according to Equation (1). We used three fixed depths, plus mixed depths to add a depth component to the task. In the fixed depth conditions, all targets were presented at the same depth (5, 7, or 9 m from the viewer). In the mixed depth conditions, the sphere at the 12 o'clock position (the top sphere) was positioned at a depth of 5 m. Each subsequent sphere in the circle (going clockwise) was 10 cm deeper than the last. See Illustration 6.
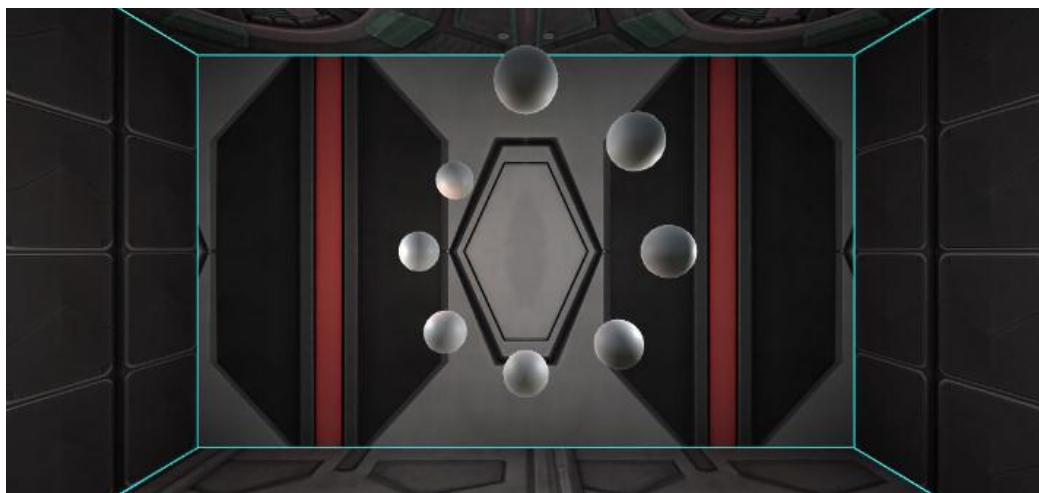
Illustration 6 Same-sized spheres in a mixed depth configuration. The spheres on the left appear smaller due to perspective, as they are farther away from the viewpoint.

All three target widths were crossed with all four depths, including mixed depth. The ordering of input method was counterbalanced according to a Latin square. There were 15 selection trials per combination of target depth and target width, hence the total number of trials was 18 participants × 3 input methods × 4 depths × 3 widths × 15 trials = 9720 trials.

The dependent variables included throughput (bits/s), movement time (ms), and error rate (%). Movement time was calculated as the time from the beginning of a trial, to the time the participant pressed the "z" key, which ended the selection trial. Error rate was calculated as the percentage of trials where the participant missed the target.

As mentioned previously in related work, we utilized Equation 2 and Equation 3 to calculate the *ID* and *TP* (throughput), where α is the rotation angle from sphere B (see Illustration 7) to sphere A and ω is the angular size the target sphere (i.e., angular interpretations of *A* and *W*). We derived the angular measures for distance and target size by trigonometric.
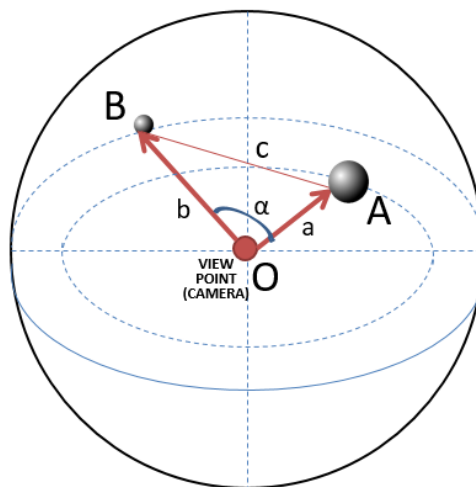
Illustration 7 The same-sized spheres A and B at different depths form triangle $\angle$AOB with the view point O. Although the straightline distance between A and B is c, the angular distance is represented by α. A similar calculation is used for the angular size of targets from the viewpoint.

Finally, we also collected subjective data via nine questions using a 7-Likert scale. These questions were based on those recommended by ISO 9241-9.

3.6 Results

3.6.1 Error Rates

We utilized ANOVA test on mean error rates for each participant. Mean error rates are summarized in Illustration 8. There was a significant main effect of input method on error rate ($F_{2,14} = 13.998$, $p < .05$). The Scheffé post-hoc test revealed that the difference between all three input methods was significant ($p < .05$). Eye-only and eye+head had much higher errors than head at roughly 40% and 30% vs. 8% respectively. The high standard deviation reveals great variation in performance, especially for eye-only and eye+head. This suggests that participants had much greater difficulty selecting targets with the eye tracker, consistent with previous results [57].
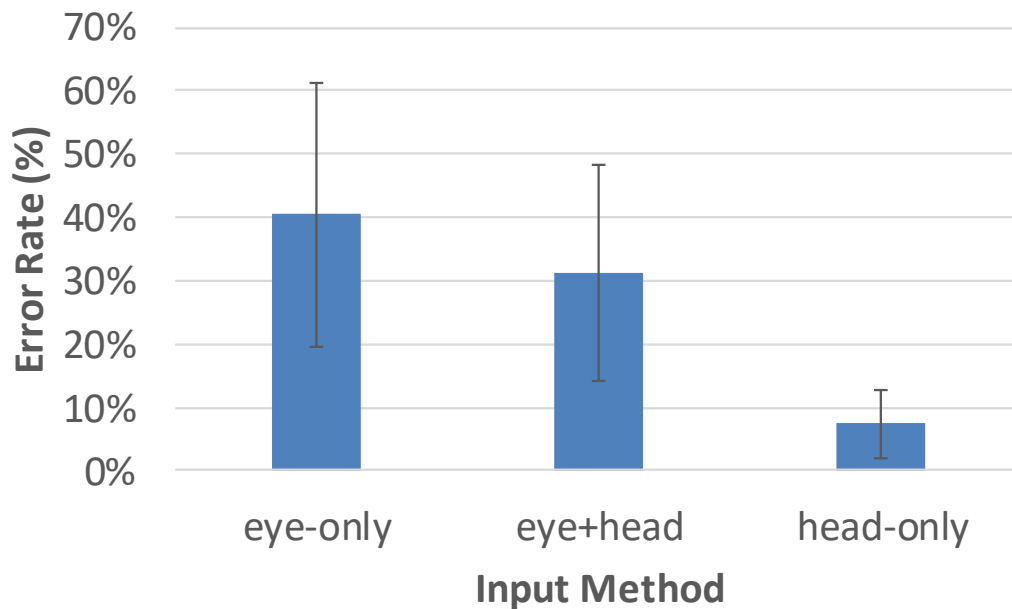
Illustration 8 Mean error rates for each input method. Error bars show ±1 SD.

Illustration 9 depicts error rates by target depth and size for each input method. Note that error rate increased for both smaller targets, and targets farther from the

viewpoint because the effect of farther targets and smaller targets are equivalent due to perspective The error rates of eye-only and eye+head increased sharply, while error rates of head-only increased only slightly. Eye-only and eye+head had a varied greatly depending on the target size and depth. Eye-only and eye+head were notably worse with the deepest target depth (9 m). The effect of target size – expected in accordance with Fitts' law – was also quite pronounced with mixed-depth targets.
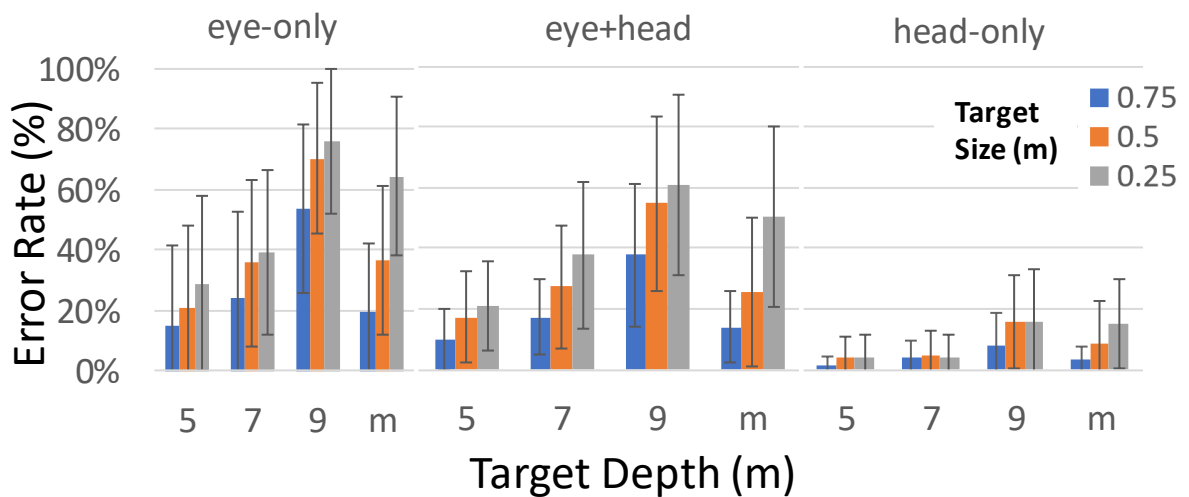
Illustration 9 Error rate by target size and target depth for each input method. Note 'm' depth represents mixed depths. A higher depth number indicates a farther/deeper target. Error bars show ±1 SD.

We note that the angular size of the target combines both target depth and size, both factors which influence error rates, as seen below. Due to perspective, a farther target will yield a smaller angular size, and according to Fitts' law, should be a more difficult target to select. Hence, we also analyzed error rates by angular size of the targets. As expected, angular size had a dramatic effect on selection accuracy. As seen in Illustration 10, we detected a threshold of about 3°. Targets smaller than this (either due to presented size, depth, or their combination) are considerably more difficult to select with all input methods studied – but especially for the eye-only and eye+head input

methods. We thus suggest ensuring that selection targets are at least 3° in size, to maximize accuracy.
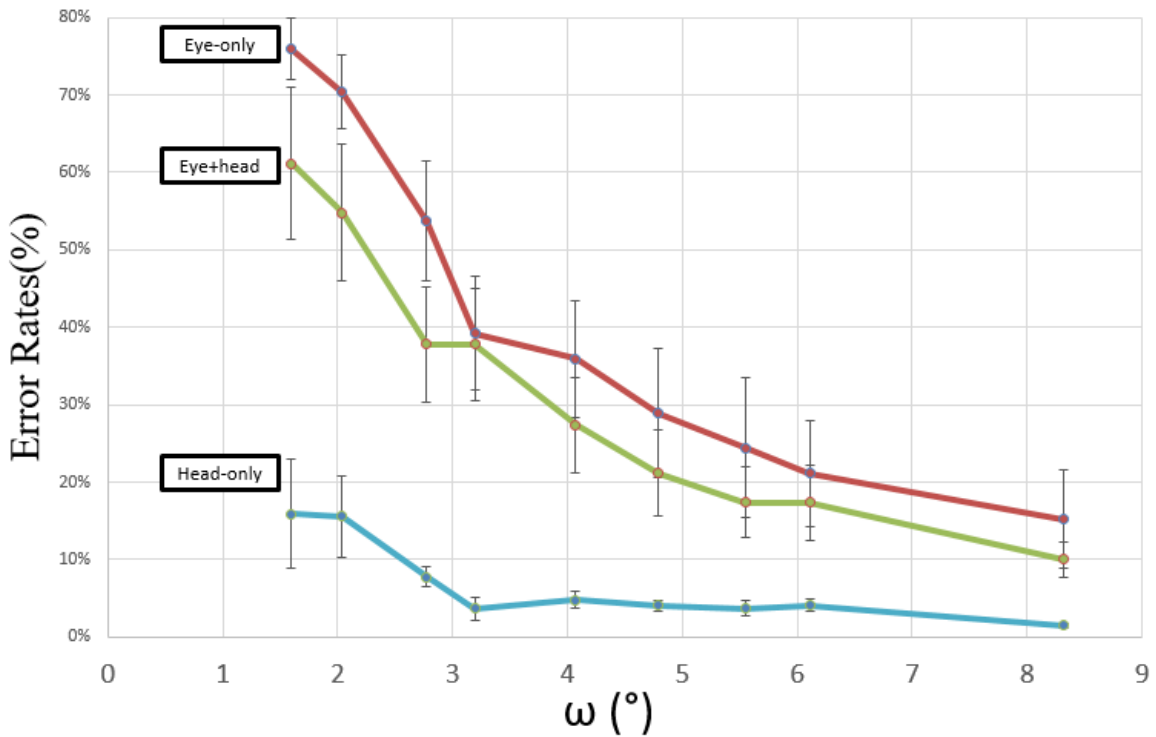
Illustration 10 Average error rate for each input method vs. angular size of the target (ω), in degrees.

3.6.2 Movement Time

We utilized ANOVA test on mean movement time for each participant. Mean movement times are summarized in Illustration 11. There was a significant main effect of input method on the movement time ($F_{2,14} = 4.713$, $p < .05$). The Scheffé post-hoc test revealed significant differences between head-only and the other two input methods ($p < .05$). Eye+head and eye-only were not significantly different from each other. This again suggests that the presence of eye tracking yielded worse performance – the one input method that did not use it (head-only) was significantly faster than both input methods that did.
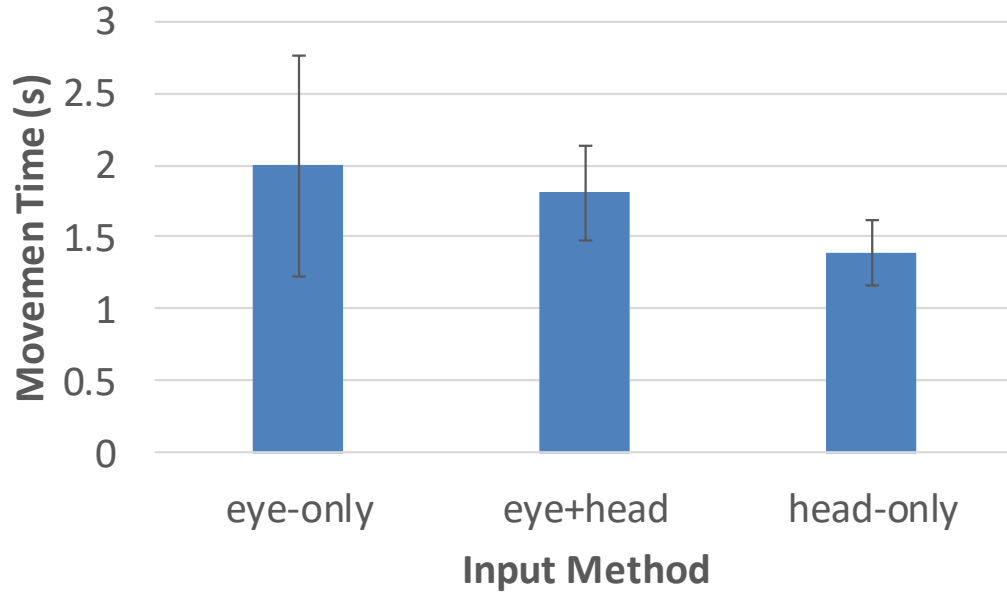
Illustration 11 Movement time by selection method. Error bars show ±1 SD.

As seen in Illustration 12, movement time increased slightly as the target size became smaller. However, the effect of target depth was more pronounced, particularly with the eye-only input method. The other two input methods increased slightly and similarly.
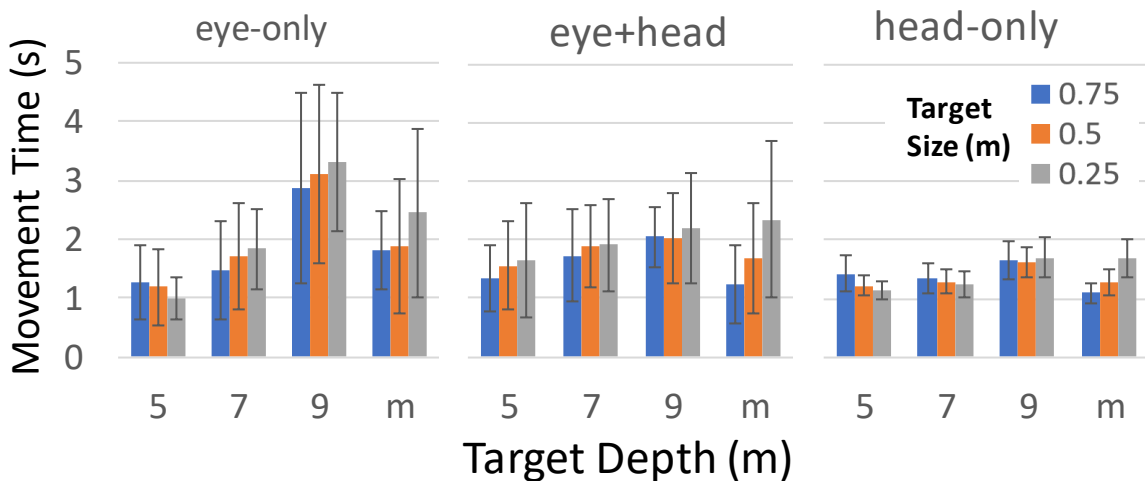
Illustration 12 Movement time by target size and depth for each selection method. Note 'm' depth represents mixed depths. Error bars show ±1 SD.

The angular size of the target also influences the movement time, as seen below. As expected like error rates, angular size had a similar effect on selection speed, especially for eye-only input method. As seen in Illustration 13, we detected the same threshold of about 3°, which were obviously shown by eye-only and head-only. All input methods took more time to select when targets became smaller (either due to presented size, depth, or their combination. There was another notable fact when the angular size was greater than 3°. When the angular size was greater, the movement times of three input methods were very close.
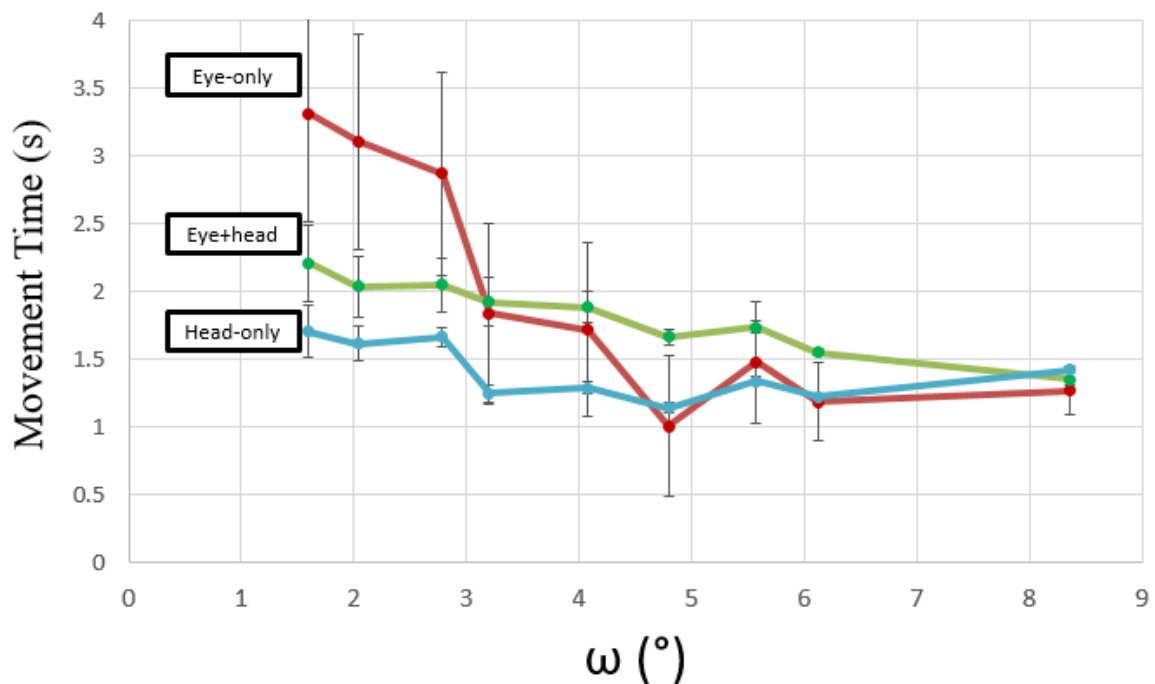
Illustration 13 Average movement time for each input method vs. angular size of the target (ω), in degrees.

3.6.3   Throughput and Fitts' Law Analysis

Throughput scores are summarized in Illustration 14. There was a significant main effect for input method on throughput ($F_{2,14} = 21.99$, $p < .05$). The Scheffé post-hoc test also showed significant differences ($p < .05$) between eye+head and head-only, and

head-only and eye-only. However, eye+head and eye-only were not significantly different, which again suggested some difference due to the presence of eye tracking. Head-only was once again the best among the three input methods. The throughput scores of eye-only and eye+head were in the range reported by Mackenzie [7], yet notably lower than average throughput for the mouse [19]. We note that throughput was also somewhat higher than that reported by Teather and Stuerzlinger [23, 59] for a handheld ray-based selection technique.
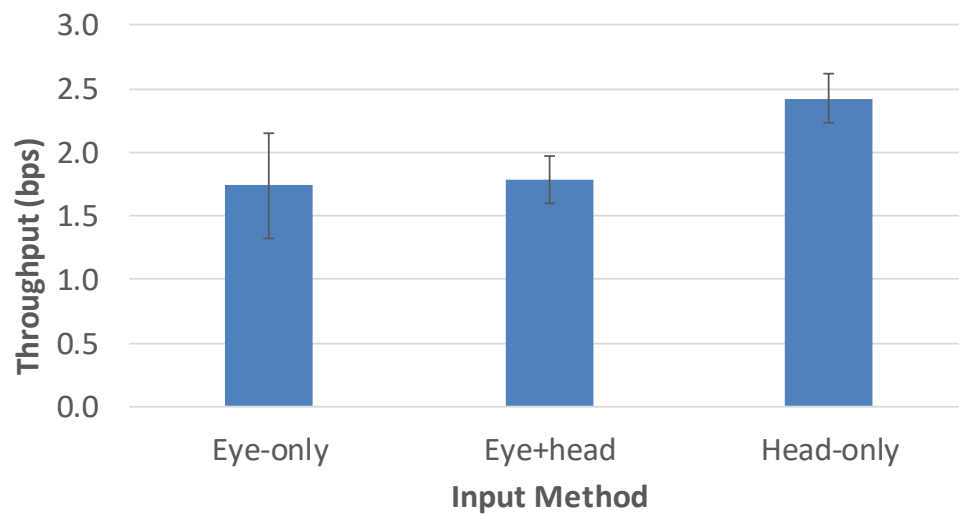
Illustration 14 Throughput by input methods. Error bars show ±1 SD.

As is common practice in Fitts' law experiments, we produced linear regression models for each selection method showing the relationship between *ID* and *MT*. These are shown in Illustration 15.
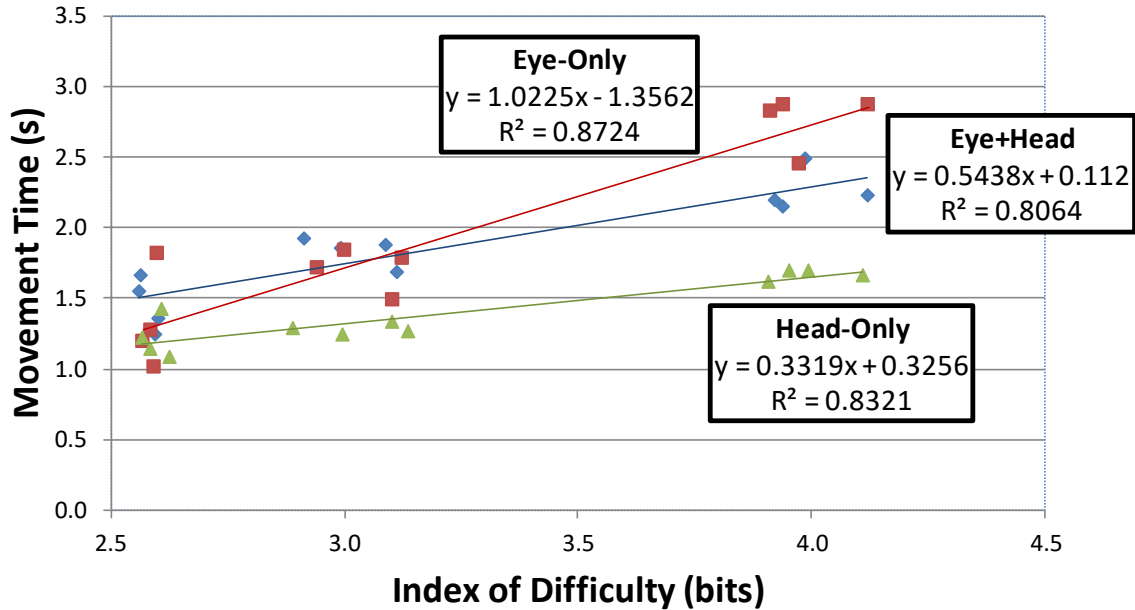
Illustration 15 Regression models for all input methods.

Note that the presented R^2 scores are quite high, ranging from between 0.8 and 0.87. This suggests a fairly strong predictive relationship between *ID* and *MT*, which is typical of interaction techniques that conform to Fitts' law. We note that these scores are somewhat lower than in other research using input devices like the mouse [19], but in line with previous research on 3D selection [23, 59]. Interestingly, the eye-only input method offered the best fitting model, suggesting that eye-tracking conforms to Fitts' law better than head-based selection [7, 56].

3.6.4   Subjective Questionnaire

The device assessment questionnaire consisted of 9 items, modelled after those suggested by ISO 9241-9. We asked each question for each input method. Each response was rated on a 7- point scale, with 7 as the most favourable response and 1 the least favourable response. Responses are seen in Illustration 16.

Overall, participants rated head-only best on all points expect neck fatigue. Eye-only was rated best on neck fatigue. Conversely, and perhaps unsurprisingly, head-only was rated best on eye fatigue, and eye-only was rated worst. Participants were also aware of the difference in accuracy offered by each input method; they reported head-only was most accurate, followed by eye+head, with eye-only rated worst, much like the error rate results shown earlier.
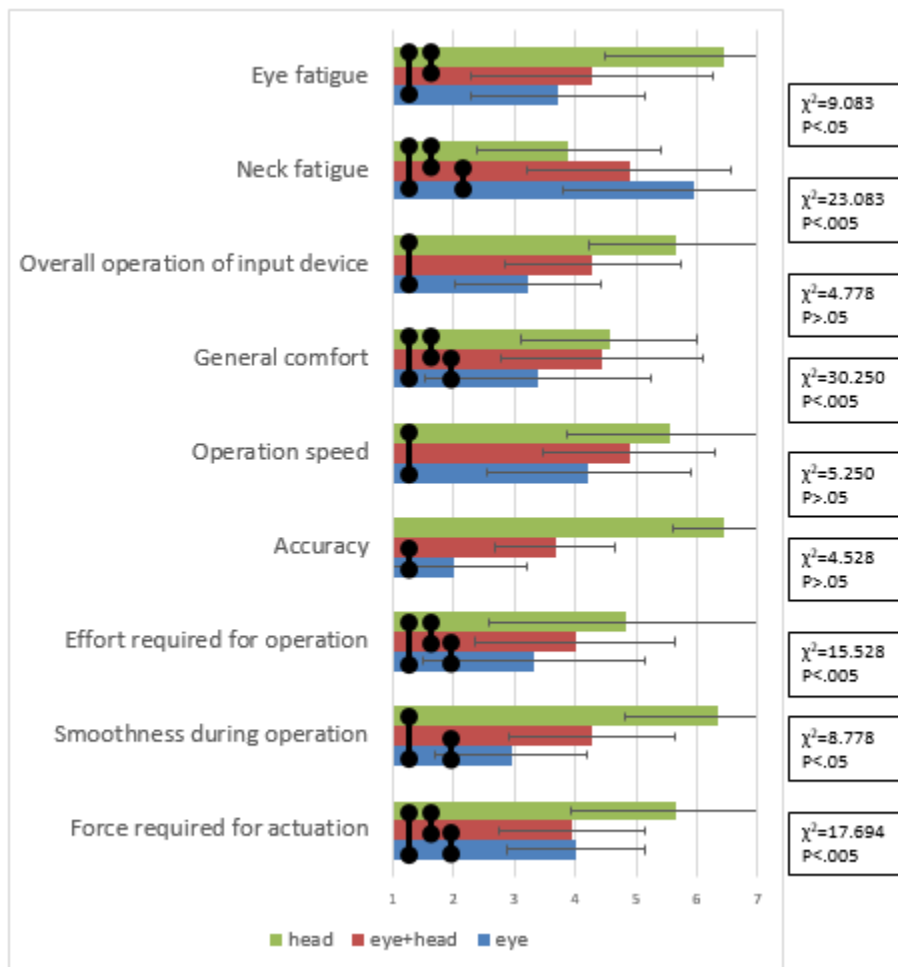
Illustration 16 Average of response scores for each survey question. Error bars show ±1 SD. Higher scores are more favorable in all cases. Statistical results via the Friedman test shown to the right. Vertical bars (⊶) show pairwise significant differences per Conover's F test posthoc at the p <0.05 level.

3.6.5   Interview

Following completion of the experiment, we debriefed the participants in a brief interview to solicit their qualitative assessment of the input methods. Eleven participants preferred head-only because it provided high accuracy, and it was the most responsive and comfortable. Six participants found eye-only the worst, reporting that it was difficult to use. Some indicated that due to their prior experience wearing standard HMDs, they were already used to head-based interaction, which may help explain their preference towards head-only. However, they tended to indicate that they found eye-only inefficient.

Five participants found eye+head the worst. Much like our initial hypothesis, at the onset of the experiment, these participants expected eye+head would offer better performance, but were surprised to find that it did not. A few participants indicated that they experienced some nausea and neck fatigue with eye+head. Finally, five participants rated eye-only the best. Although it did not provide accurate operation, these participants felt comfortable using it. They also complained about neck fatigue with both head-based input methods, and indicated that they looked forward to wider availability of eye-tracking HMDs in the future. Some even suggested that for tasks that did not require precision, they would always choose eye-tracking.

3.7   Discussion

Before the experiment, we hypothesized that using eye and head tracking together – the eye+head input method – would offer the best performance of the three input methods, since it offered the best capabilities of both eye- and head-tracking. Our data, however, disproved this hypothesis. In fact, the head-only input method performed the best across all dependent variables, especially accuracy. In contrast, the two input

methods utilizing eye tracking (eye-only and eye+head) were fairly close in performance, with eye+head generally performing better than eye-only. We hypothesized that head-only would yield the lowest error rates; this hypothesis was confirmed. We also hypothesized that participants would prefer eye+head, but this was not the case. Based on past work, we had expected that eye+head would provide a selection method consistent with how we use our eyes and head together in pursuit tracking [37]. However, during testing, we observed that the cursor sometimes jittered, resulting in poor precision with eye+head. This may be a limitation of the hardware.

Previous eye tracking research relates the importance of calibration problems, which can drastically influence the data [6, 37, 38]. Two potential participants were excluded because despite 5 attempts, they still failed the calibration. This might be an inherent flaw of FOVE's calibration algorithm or hardware. We also observed that calibration quality greatly influenced selection performance. For example, during the calibration phase, participants had to follow a moving green dot with their eye gaze. One participant mentioned that the green dot stopped moving for more than 3 seconds on the 9 o'clock and 1 o'clock direction. This may be due to a software bug, or because the eye tracker had difficulty detecting the participant's eyes. As a result, during testing, that participant could not reliably select targets in those directions, necessitating re-calibration of the eye tracker. In all these sessions, although the participant had passed the calibration component, such pauses during the calibration process could still yield poor results, likely affecting performance with both eye-tracking input methods. Participants suggested improving the calibration process in future, which may yield better results with the eye-only and eye+head input methods.

As detailed above, participants strongly favoured the head-only input method. In the eye-only and eye+head sessions, participants indicated that they could comfortably and reliably select larger spheres. However, when spheres were smaller and/or deeper into the scene (i.e., smaller in terms of angular size), participants felt very frustrated and uncomfortable, particularly when missing the targets. Based on this observation, and our earlier analysis of angular sizes, we recommend designers to avoid targets smaller than c in size. While it is well-known that target size influences pointing difficulty [19, 74], this seems especially important with eye-only selection. In contrast, large targets and relatively closer targets are considerably easier for participants to select. Thus, future work would investigate the selection performance of eye-based techniques on much larger targets. On the other hand, the eye's microsaccade amplitudes vary from 2 to 120 arcminutes [75, 76], or a maximum of about 2°. Meanwhile, the FOVE HMD has around or less than 1° tracking precision. The total c deviation meets our testing result. Based on MacKenzie's error rate results [7] that targets in 32 and 16 pixels (1.15° and 0.57°) by dwell had error rates of 55% and 84%, while blink had lower error rates of 38% and 55% because it avoided microsaccade, thus our 3° threshold with around 40% error rate is acceptable.

Interestingly, during the interview most (16/18) participants felt that eye+head would work well in VR first-person shooter games, despite the largely negative results yielded by this input method. Participants suggested that head-only could cause sickness, and eye-only is too inaccurate. Participants suggested that an improved version of eye+head would work well for shooting. Similarly, half felt eye-only would work well

for menu selection, while the rest thought head-only would work best. One suggested that assuming large enough widgets, any technique would be effective.

4   Chapter: Travel Study

Our selection study has shown that eye-tracking tends to offer poor performance in 3D selection tasks and could only perform well on certain targets. However, a few researchers [26, 27] have already been interested in eye-based travel that might be benefited from eye's smooth pursuit. In general, travel has a lower accuracy requirement than selection. Travel works well enough if users can get into the general vicinity of where they intended to get. So in light of this lowered accuracy requirement, it is reasonable that eye-based input may work better for travel than selection. Thus, in our second study, we presented a study comparing the performance potential of eye-tracking as an alternative to other travel control techniques.

We developed a travel testbed virtual environment where the user flies through ring to compare gaze-directed steering using the eye to that with the head. To provide a baseline of comparison, we also included mouse and joystick-based steering. Our study thus included 7 input techniques to control the flying direction. Four were "single input" and three were "combination input". The single input techniques controlled both the head orientation and movement direction simultaneously, similar to first-person shooter game controls. These included: 1) head-only, 2) eye-only, 3) mouse-only, 4) joystick-only. With the exception of head-only, head-tracking was disabled in these input techniques. In the combination input methods, head-tracking was enabled in tandem with three other input techniques to control movement direction. They were: 5) head+eye, 6) head+mouse, 7) head+joystick. The mouse-only input technique was the baseline of this study. While mouse-based steering is atypical in VR travel, it is very common in first-person shooter games (used in tandem with keys to control movement speed). The

joystick is a traditional gaming controller, we tested it with the purpose to see the performance of this widely used representative in VR traveling.

4.1 Hypotheses

Our hypotheses included:

H1: Among the 4 single input techniques, mouse-only should have the fastest completion times, while the joystick would take the longest time.

H2: Among the 4 single input techniques, eye-only would perform better than head-only because it reduces the head movement time, but it might fly farther to the central ring than mouse and head because of eye's inherent inaccuracy.

H3: Among the 4 single input techniques, head-only would cause the lowest cyber sickness because the consistency of visual information and vestibular system.

H4: Among the 3 combination input techniques, head+mouse is still the baseline, it should have the best performance. The head+eye would be better than head+joystick because it's more intuitive for steering.

H5: Among the 3 combination input techniques, participants should prefer head+eye the most because the other two combination input techniques need extra body movements (operating mouse and joystick).

H6: All the combination input techniques should have the shorter completion time and accuracy than their corresponding single techniques.

4.2 Participants

We recruited fourteen participants (aged 18 to 40, $\mu = 27$ years, 8 male). All were daily computer users ($\mu = 5$ hours/day). Five had prior experience with eye tracking. Three had no prior VR experience, another three had limited VR experience (having used

it once or twice ever), and the rest used VR on average around 5 times per month. All participants had colour vision. Five had normal vision, while the rest had corrected vision. All participants could see stereo, as assessed by pre-test trials. All participants were very familiar with games, 4 were frequently video game users (μ=5 times/week). One potential participant could not pass the calibration and two potential participants withdrew from the pre-test trials due to the nausea.

4.3   Apparatus

The study was conducted using a VR-capable laptop with an Intel Core i7-7700HQ CPU, an NVIDIA GeForce GTX 1070 GPU, and 16GB RAM. Participants wore a FOVE VR HMD during testing. The FOVE has a display resolution of 2560 x 1440 with a 100° field of view. It offers IMU-based sensing of head orientation, and optical tracking of head position, but does not provide IPD correction. The FOVE includes two integrated infrared eye-trackers that offer tracking precision of less than 1° at a 120 Hz sampling rate. We also utilized a wired mouse and an Xbox controller as other input devices.

We developed the experimental interface and testbed using Unity 5.5 and C#. The experiment included a typical flying task; to this end, the software presented three sets of rings in the air with the simple background of the blue sky over a desert and lake terrain. Participants were tasked with flying through these rings using the current control scheme. See Illustration 17. The desert terrain was the reference object that enabled participants to feel the relative speed of motion. All the tasks were conducted in the air, no collisions occurred with the terrain.
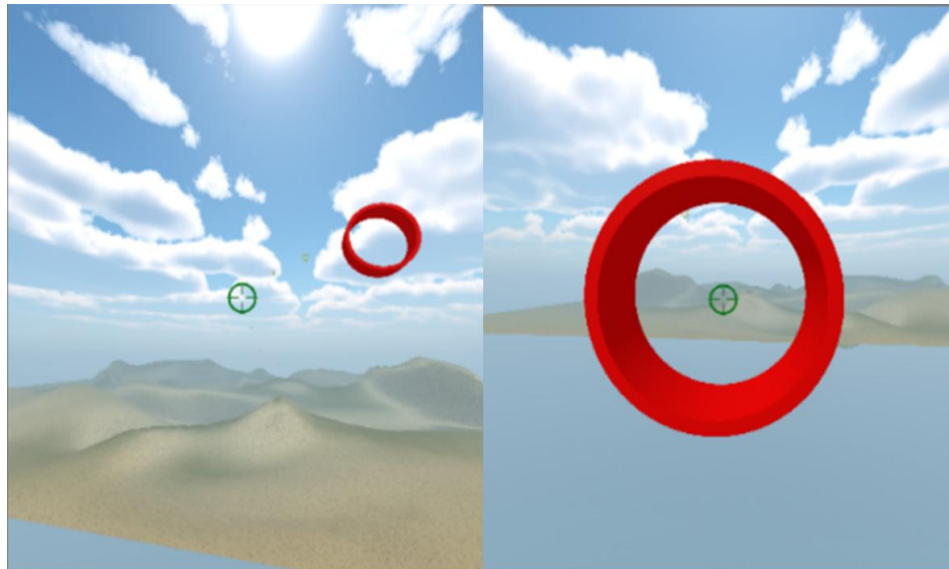
Illustration 17 Experimental task showing the terrain, skybox, and rings the participants flew through.

The experimental task presented eight yellow ring in spiral arrangement in the air. The target ring was highlighted red. See Illustration 18. Depending on the condition, the rings were put in 10º, 20º, or 30º deviations with respect to the previously passed ring. The distance (z-axis) between each ring was all 100 meters. The radius of each ring was 1.5 meters. The width of each ring was 1 meter. The 1-meter width ensured the software could reliably detect the collision point (in the plane of the ring) when the participant passed through the ring. The software started to scan the inside of the ring when the view camera hit the surface plane of the ring. The frame rate in the experiment was stable at 80 fps. In order to reduce these effects, we utilized a fixed linear speed level. We tested several velocities in the pilot study and finally chose the default value in Unity as this appropriate speed only when the user could not feel a significant sickness when tilting and rotating. The software also displayed a green round cursor to facilitate steering towards the targets (see Illustration 17). The input devices were operated by effectively controlling a cursor that defined the orientation of the movement vector, originating at the

head. We recorded all the coordinates of the collision points with the plane of each ring (to facilitate accuracy measures, i.e., distance from the ring centre), including inside and outside the ring, the successes and failures.
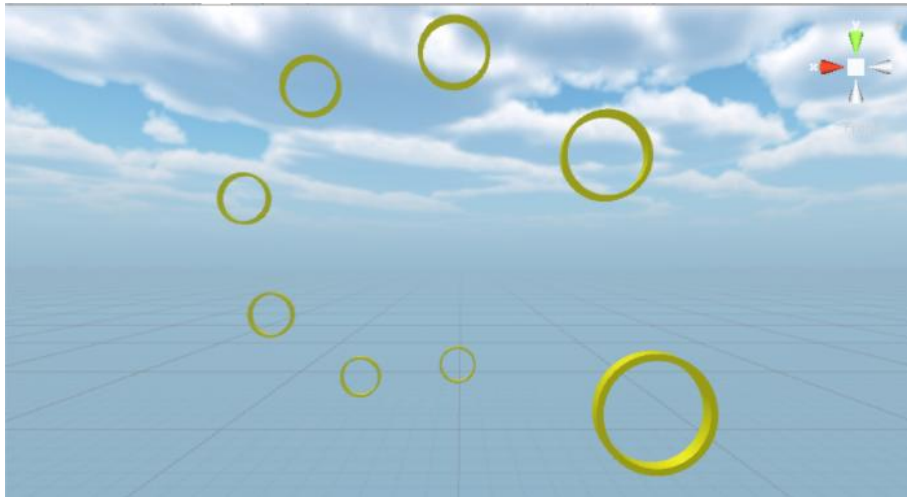
Illustration 18 The ring arrangement: the rings were put in 20-degree deviation in one block

4.4   Procedure

Upon arrival, we first briefed participants on the motivation, goals, and procedure for the experiment, then provided them with consent forms and demographics questionnaires. Then we provided a demo video of the interface and introduced them how to operate each of the travel techniques. All participants first completed the FOVE calibration process, which took approximately one minute. Calibration involved gazing at a green dot that appeared at a circular position on the display. We also used this calibration process as pre-screening for the participants: Potential participants who could not complete the calibration process cannot take part in the experiment. Prior to each new session using the eye tracker (i.e., eye-only and head+eye), the eye tracker was re-calibrated to ensure accuracy throughout the experiment. Because the eye-only and

head+eye were the only novel techniques compared by other input techniques, and all participants had prior experience with the mouse and joysticks, many (8/14) participants even indicated that they were very familiar with the use of head-based orientation in VR, we added a few extra time for practice trials in these two techniques. We started the actual sessions before they thought they were managed to control their eyes properly.

Participants were instructed to fly through the red highlighted ring using the current travel technique. They were instructed to fly as closely as possible to the center of ring. When the participant started the test, all of rings appeared in front of the view and the first target ring was in red color. Because of the distance between rings, participants were not able to see all rings, but could see the next three or four rings in the view. As participants travelled through the rings one by one, the remaining rings appeared. Upon passing each red (target) ring, it would disappear and the next ring in the sequence would turn red.

A block had 8 rings, each representing a different trial, and each in one of 8 different directions, organized in a spiral/corkscrew configuration. See Illustration 18. Each travel technique testing session consisted of 3 such blocks. An extra "practice" ring was added to each block to help participants get used to a new condition. Data for this practice ring was excluded from our analysis. Regardless if the participant flew through, or missed (outside) the target ring, the next ring would highlight red. If they flew outside the ring, the trial was recorded as a miss.

Upon completing a session, participants completed three questionnaires, the NASA-TLX, the SSQ, and the traveling performance questionnaire developed by Bowman et al. [26]. Upon completing all seven sessions, participants completed a 5-point

questionnaire of their preference of the seven input techniques and were debriefed in a short interview. Our experiment took approximately 70 minutes in total for each participant, for which they were compensated $10.

4.5 Design

The experiment employed a 7 × 3 within-subjects design. The independent variables and their levels were as follows:

Travel technique: Eye-only, head-only, mouse-only, joystick- only, head+eye, head+mouse, head+joystick

Difficulty: 10°, 20°, 30°

Since we considered each ring a single trial, in total, each participated completed 7 × 3 × 8 × 3 = 504 trials. Across all 14 participants, this yielded 7056 trials. Difficult was represented as eccentricity of the next ring (i.e., necessitating a 10°, 20°, or 30° rotation from the previous ring). Difficulty was arranged from the easiest to the hardest, i.e., the first three blocks were 10°deviations, the second three blocks were 20° deviations, the last three blocks were 30° deviations. Ordering of travel technique was counterbalanced according to a Latin square.

The dependent variables were completion time, success rates, collision radius, traveling performance, NASA-TLX and SSQ. The completion time was the entire time to complete three blocks on the same difficulty level. The success rates are the percentage of rings successfully passed per difficulty in each session. The collision radius represented all the collision points in the circular range of 3m from the centre of the ring, the successful and failed points within 3m were all included.

4.6   Results

4.6.1   Completion time

Mean completion times are summarized across the travel techniques and three difficulty levels in Illustration 19. There was a significant main effect of travel technique on the completion time ($F_{6,273} = 38.607$, $p < .001$), and no significant effect of difficulty on the completion time ($F_{2,273} = 0.96$, ns). The travel techniques and three difficulty levels had no interaction effects with each other. ($F_{12,273} = 0.248$, ns). Overall, participants tended not to take much longer regardless of difficulty. The reason might be that degree deviations were not great enough to differentiate with each other. We utilized the Tukey-Kramer post-hoc test to detect pair-wise differences between travel techniques. Both joystick techniques yielded much worse performance than others. The head did not help improve the speed.
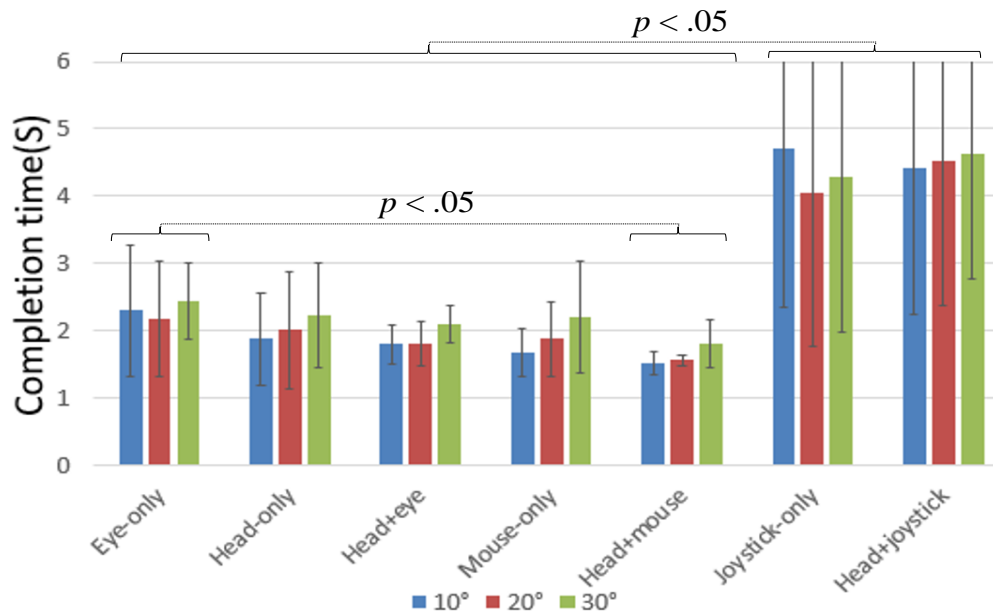
Illustration 19 Mean completion time by travel techniques on three difficulty levels. Error bars show ±1 SD. Braces and dashed lines indicate "clusters" of travel techniques that show pairwise significant differences via post-hoc testing at the $p < .05$ level.

4.6.2 Success rates

Illustration 20 depicts success rates by difficulty levels for each travel technique. There was a significant main effect of travel technique on the success rates ($F_{6,273} = 20.41$, $p < .001$), and no significant effect of difficulty on success rates ($F_{2,273} = 1.449$, ns). The travel techniques and three difficulty levels had no interaction effects with each other. ($F_{12,273} = 0.232$, ns). The Tukey-Kramer post-hoc test showed pair-wise differences ($p < .05$) between travel techniques.

Illustration 20 Mean success rates by travel techniques on three difficulty levels. Error bars show ±1 SD. Braces and dashed lines indicate "clusters" of travel techniques that show pairwise significant differences via post-hoc testing at the $p < .05$ level.

4.6.3   Coordinate Map and Collision Radius

Illustration 21 shows coordinate maps for each travel techniques of all degree levels cut from the z-axis plane of all collisions within a 3 m radius of the ring. The red circle indicates the target ring with 1.5 m radius. This visualization gives a good indication of the degree of control offered by each of the travel techniques; conditions more closely clustered near the center of the red circle indicate participants were better able to stay near the ring centre while traveling. Conversely, conditions with many data points outside the circle indicate travel techniques where participants had greater difficulty.

We note that joystick-only is much sparser than the other travel techniques. This is because there were also collisions that happened farther than 3 m, but we exclude these from Illustration 21 for space considerations. Mouse-only offered consistently high precision, hitting virtually all the collisions within the ring. Head+eye was a bit sparser than head-only, but both of these did well overall. Eye-only, joystick-only and head+joystick all had many collisions out of the ring, joystick-only was the worst. This map revealed the consistency with success rates.
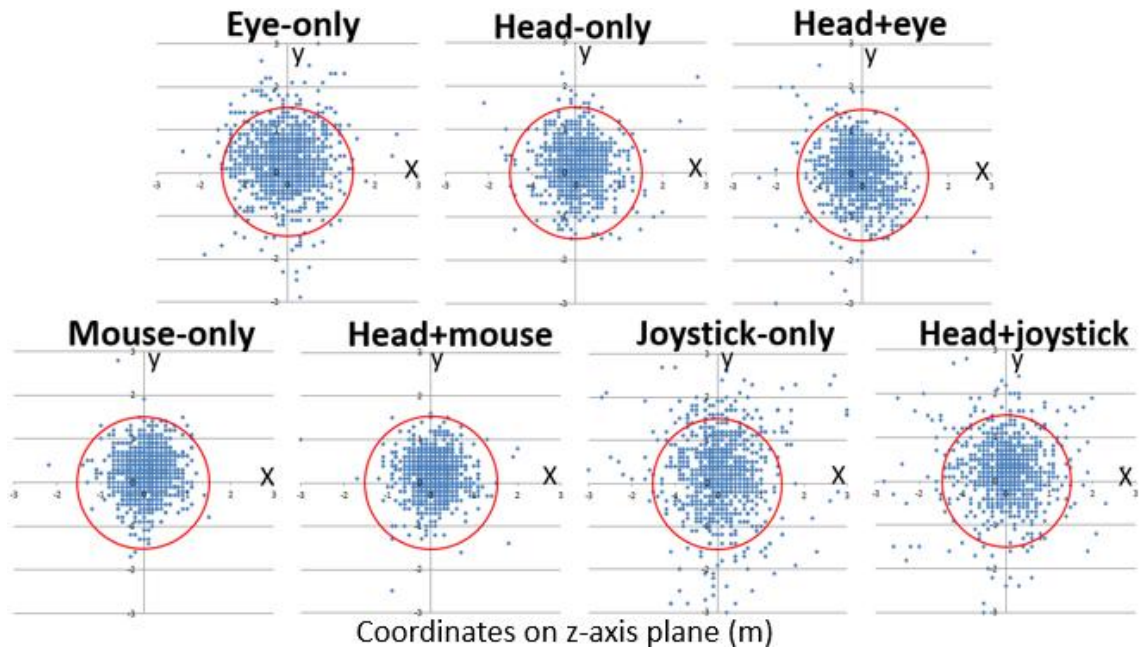
Illustration 21 Coordinate maps on z-axis plane for each travel technique, across all three difficulty levels. The red ring depicts the target ring, and each blue mark depicts a coordinate. This includes all trials for each travel technique, aggregated together.

We also analyzed the mean radius of collisions – i.e., the magnitude of error. These are summarized in Illustration 22. The radius represents how far the actual path deviated from the optimal path, where the collision point should be at the center of the ring, so the greater the radius, the less accurate the technique was. There was a significant main effect of travel technique on the collision radius ($F_{6,273} = 15.108$, $p < .001$), and no significant effect of difficulties ($F_{2,273} = 2.192$, ns). The travel techniques and three difficulty levels had no interaction effects with each other. ($F_{12,273} = 0.09$, ns). The Tukey-Kramer post-hoc test showed pair-wise differences ($p < .05$) between travel techniques.
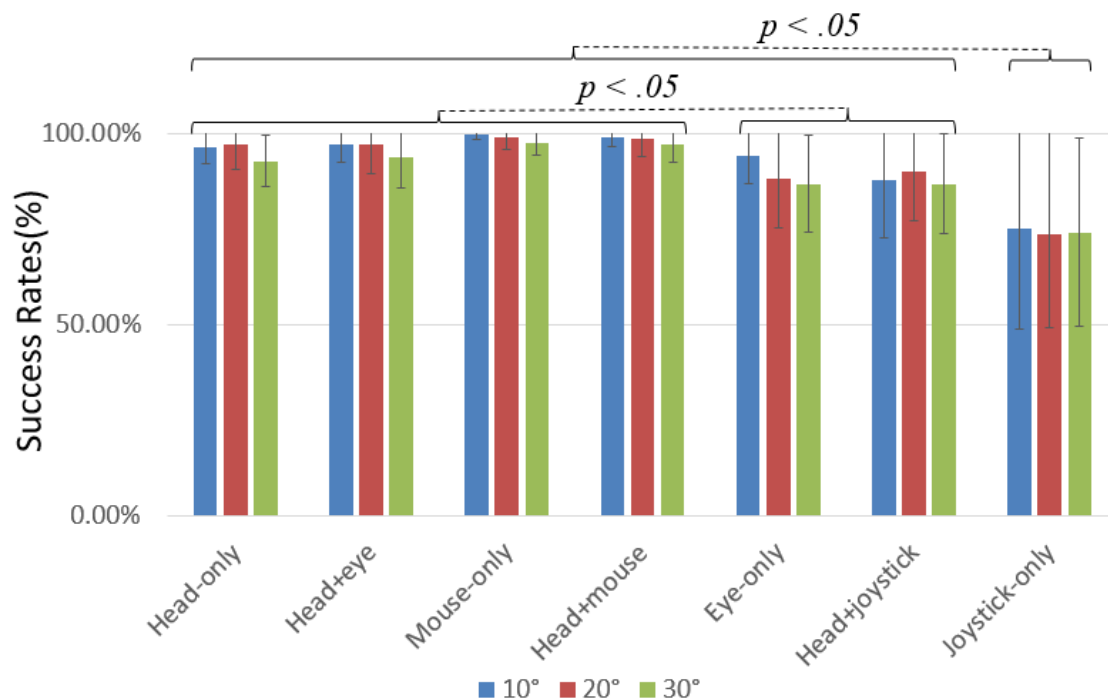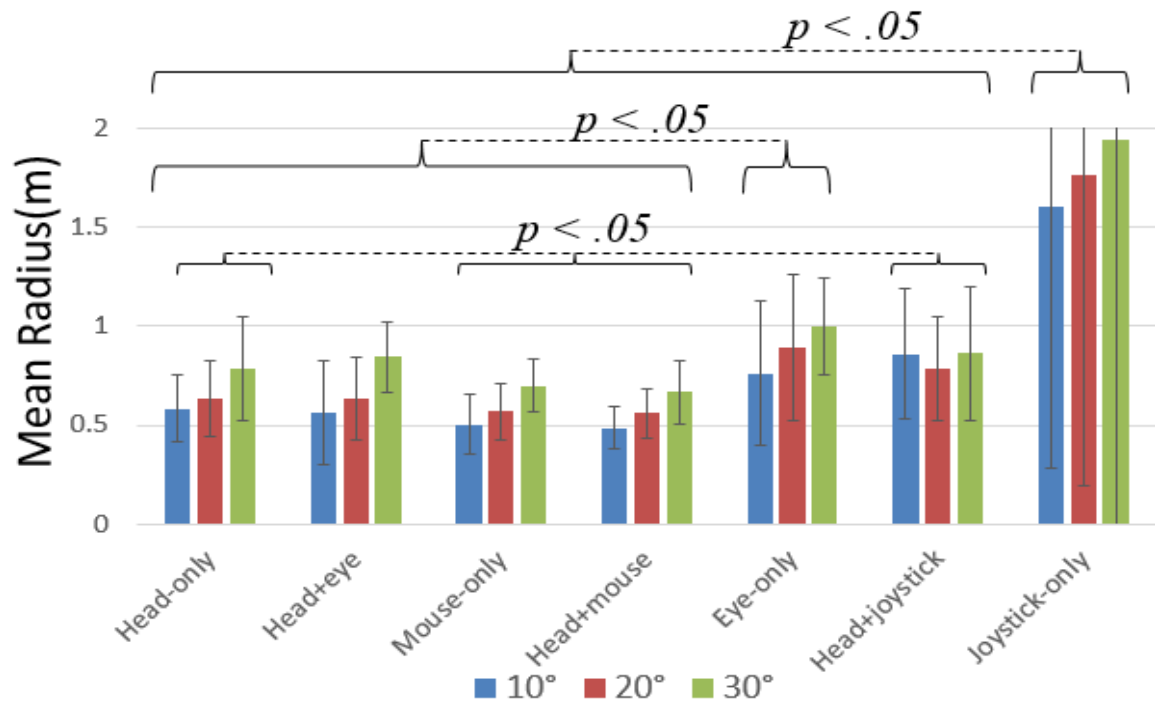
Illustration 22 Mean radius of the collision points of 10, 20 and 30-degree levels. Error bars show ±1 SD. Braces and dashed lines indicate "clusters" of travel techniques that show pairwise significant differences via post-hoc testing at the *p* < .05 level.

4.6.4   Subjective Measures

We included three questionnaires to garner subjective data on the conditions. The first was the travel performance questionnaire consisting of 5 items, and based on Bowman's travel questionnaire [26]. We asked participants to fill this questionnaire after finishing each session. Each participant rated on a 5-point scale, with 5 as the most favorable response and 1 the least favorable response. Scores from this questionnaire are summarized in Illustration 23. Overall, participants rated mouse-only and head+mouse best on all points. Head-only was rated lower than head+mouse, but still higher than eye-only and head+eye on all points. Head+eye was better than eye-only on special

awareness, while eye-only is better on the learnability. A few participants said they were confused about using the eye combination technique for the first few trials.

Illustration 23 Average of response scores for travel performance question. Error bars show ±1 SD. Higher scores are more favorable in all cases

We conducted the simulator sickness questionnaire (SSQ), based on Kennedy et al. [66]. The SSQ is commonly used to assess levels of cybersickness in virtual reality systems. The questionnaire consisted of 16 items and had 3 weighted symptom categories, i.e., nausea, oculomotor and disorientation. We asked participants to fill this questionnaire after finishing each session. Joystick-only, eye-only and head+joystick had much higher symptoms than other techniques on all three profiles. The joystick-based techniques were the worst, but eye-only also had high symptoms. The Kruskal-Wallis test revealed that there were significant differences among the seven travel techniques on general discomfort, difficulty focusing, salivation increasing, sweating and difficulty concentrating.

1	General discomfort	S
2	Fatigue	
3	Headache	
4	Eye strain	
5	Difficulty focusing	S
6	Salivation increasing	S
7	Sweating	S
8	Nausea	
9	Difficulty concentrating	S
10	Fullness of the Head	
11	Blurred vision	
12	Dizziness with eyes open	
13	Dizziness with eyes closed	
14	Vertigo	
15	Stomach awareness	
16	Burping	

Table 1 Statistic results on sixteen symptom categories among seven travel techniques via the

Kruskal Wallis test
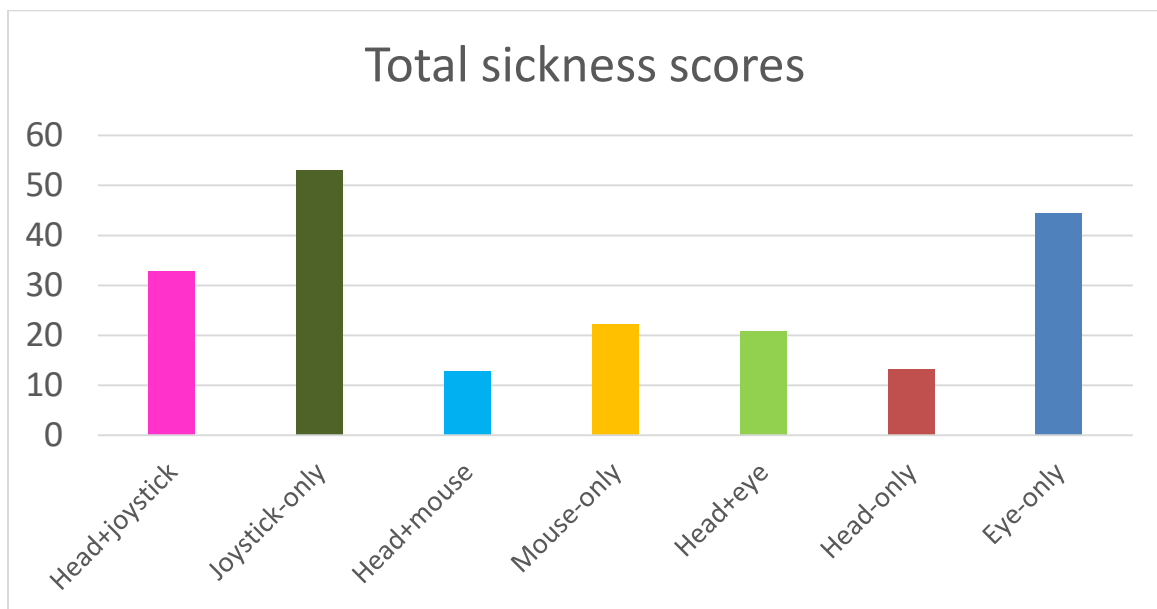
Illustration 24 Total weighed scores for SSQ question

Finally, we also used the NASA-TLX questionnaire to evaluate the workload for each travel technique. Each response was rated on a 21-point scale, with 21 as the most favourable response and 1 the least favourable response for performance, vice versa for

other 5 items. Scores are seen in Illustration 25. Unsurprisingly, and consistent with our objective performance measures, mouse-only and head+mouse were rated the highest on all scales, followed by head+eye, head+only and eye+only. The joystick techniques were rated the worst.
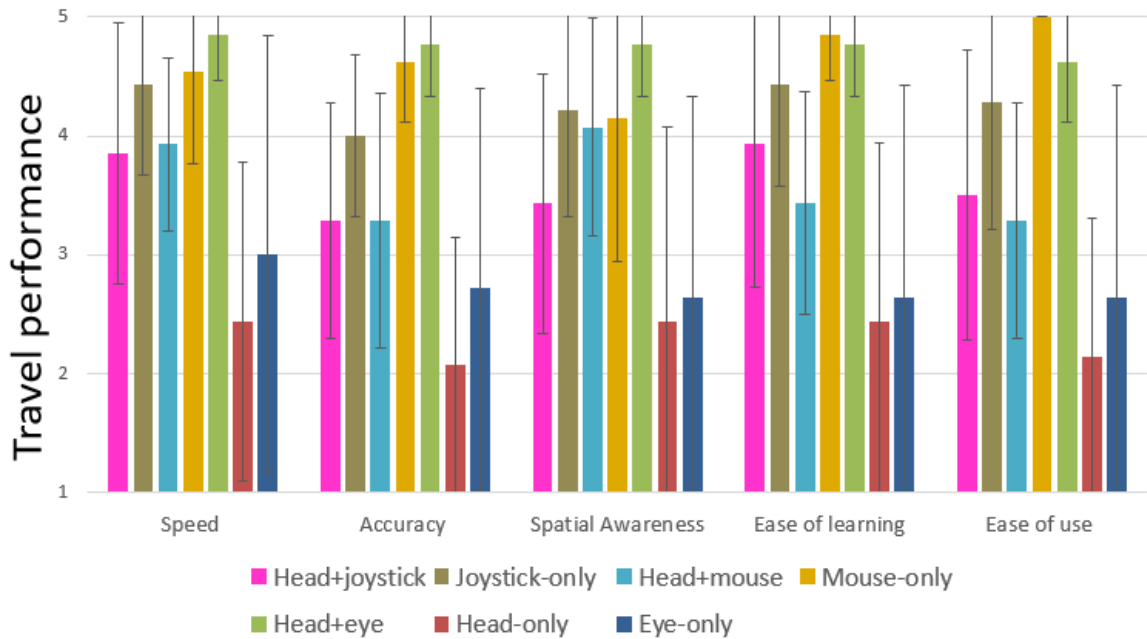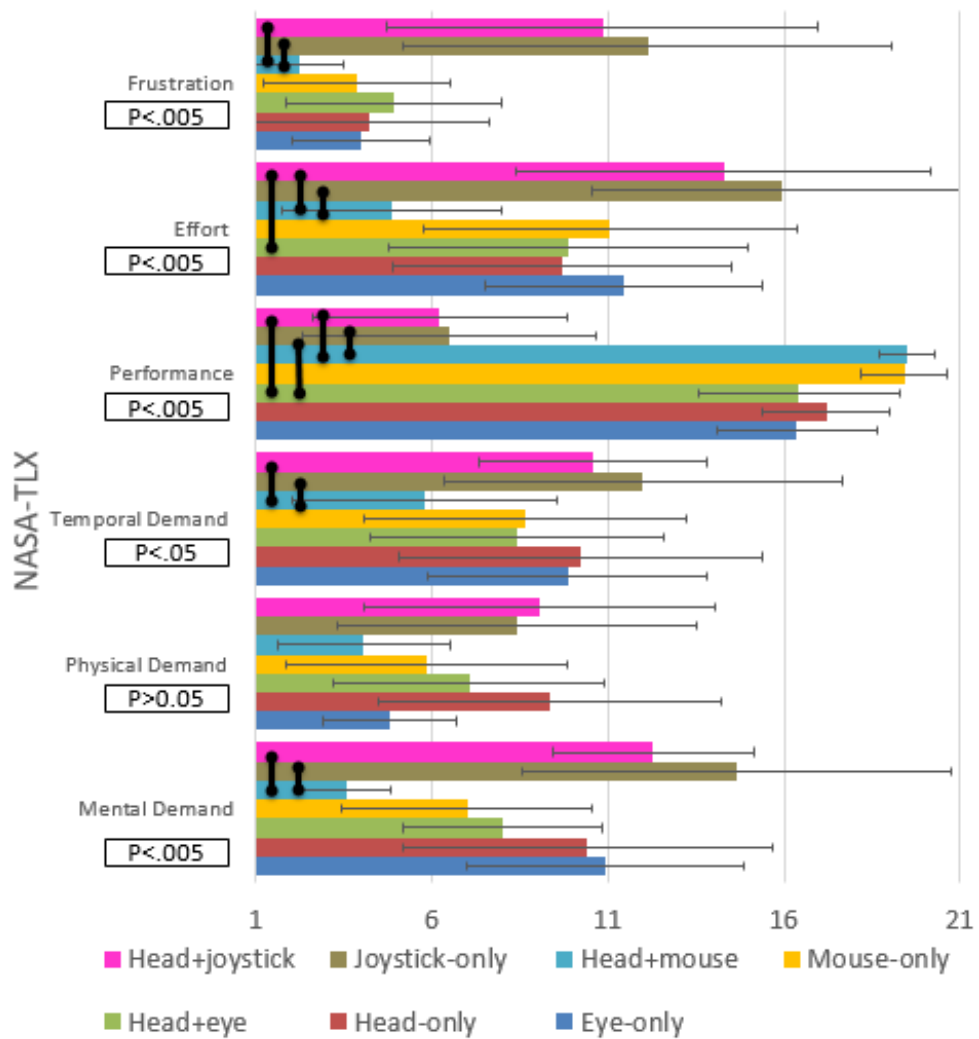

Illustration 25 Average of response scores for each NASA-Task Load Index question. Error bars show ±1 SD. Higher scores are more favorable in all cases. Statistical results via the Friedman test shown to the left. Vertical bars (⊶) show pairwise significant difference.

4.7 Discussion

After the experiment, we also conducted a short interview with participants. We asked them their preference towards each travel technique. Most participants (12/14) liked the head+mouse the most. They pointed that they used the head for prime technique and the mouse for assistance, corresponding to big movements and small corrections respectively. They said they felt the most comfortable and confident. The eye-related techniques were also mentioned by participants. Five participants rated the head+eye and three rated the eye-only as the second favorite technique followed by head+mouse. They did not select head-only because it caused much more movements than eyes. Participants noted that the head+eye or eye-only techniques could also be used for hands-free interaction, e.g., used with a combination of voice commands, blinks, or dwells.

In terms of eye-based techniques, calibration and leaning effects extremely influenced the performance. Most participants (13/14) had never used eye-tracking in VR, all of them experienced certain level of learning and adaptation depending on individual differences. The eye-only and head+eye were the only novel techniques compared by other input techniques. Therefore, we added a few extra time for practice trials for eye-based techniques. We started the actual sessions before they thought they were managed to control their eyes properly. Most participants adapted to eye-control in around a minute of practice, but some took slightly longer. However, we found the extra few minutes' trials would not be sufficient for this novel technique, suggesting the need for a future longer-term study. A few participants also suggested that more training would enhance the eye performance. Unfortunately, because in the limitation of the entire experiment time, we did not provide them more training trials than 4 minutes, including

each time for calibration. However, these conclusions were all based on participants' subjective perspectives and our observations. Thus, we highly recommend that a longitudinal study would reveal more information about how long and how well the user performance would be improved significantly after trainings.

Contrary to our expectation, our results indicated that the angular difficulty levels did not yield substantial performance differences for most techniques, particularly, eye-only and head+eye. We observed most of participants flew much more accurate and smoothly at the end of the session than the first few trials. It is possible that more extreme eccentricity angles would yield higher difficulty travel tasks. On the other hand, we did not counterbalance the difficulty levels, which might have improved the results affected by learning effects, is a limitation of this study.

As for the calibration, one potential participant could not pass the calibration after more than 5 attempts. Two potential participants passed the calibration but could not control their eyes properly, i.e., they lost the orientation after calibration and could not focus on the target ring by their eyes. We tried to recalibrate for five times but they still could not control the cursor. This yielded a great degree of jitter, which in turn caused a moderate level of nausea. We thus stopped the trials for these participants and they withdrew from the experiment. Other participants also felt certain level of nausea in the first few trials or in the middle of the session when inaccuracy occurred. This likely contributed to the higher SSQ levels with the eye-only travel technique. In the sessions, we found if the participants did not tie the HMD belt very tightly, the relative distance would be changed after moving the head thus caused the inaccuracy. Most of participants could notice the accurate difference with the first trials. Then, we asked them to

recalibrate and restart the session. Consequently, the calibration mechanism, HMD weight and the design of belts all extremely affect the accuracy. In the head+eye session, the head likely compensated for the limits of eye calibration, the participants could adjust the move direction by moving their head slightly as long as the movement was not so strenuous to change the relative position between the HMD and eyes.

Overall, joystick-only performed the worst across all dependent variables. To maintain consistency with other techniques, we utilized the left joystick to control the view direction and thus the movement vector. Two participants with extensive gaming experience found this quite natural and comfortable, but they pointed out it was always harder to control the joystick in the air than on the ground. The head+joystick had higher standard deviations than others for completion time and success rates. The reason might be the different traveling strategies used by participants. Some participants liked to use the joystick as the dominant technique but a few of them liked to use the head as the dominant technique especially for larger degree deviations between rings.

In reviewing our hypotheses, we confirmed that joystick-only had the longest completion time, head-only yielded the least cyber sickness, head+eye performed better than head+joystick, the head+eye and head+joystick really improved their corresponding single input techniques on for all objective evaluations and subjective feelings. However, head+mouse performed the best, and eye-only was not better than head-only. Therefore, H1, H3, H4 and H6 were confirmed, while the rest were rejected.

5 Chapter: Conclusion

5.1 Summary

It seems likely that eye-tracking will become available in more head-mounted displays in the near future. In this thesis, we developed two different testbeds for VR selection and navigation. We implemented three input techniques for a selection study and added four more input techniques for travel study.

While eye tracking has been used previously to support selection tasks [7, 50, 51, 56], our selection study was the first to look at eye-only selection performance in a VR environment using Fitt's law and ISO 9241-9. We founded that the head-only selection offered the fastest selection times and the best accuracy. Moreover, it was strongly preferred by participants. The combination of eye-tracking and head-based selection (our eye+head input method) performed roughly between the other two, failing to leverage the benefits of each. Our results indicate that, at least for the time being and in the absence of more precise eye trackers with better calibration methods, head-only selection is likely to continue to dominate VR interaction. As for the subjective satisfaction, the participants preferred the head-only the best, while the eye-only was the worst.

In our travel study, we explored the performance of eye-based technique in VR traveling based on our flying testbed. We implemented seven travel techniques to compare the single devices and combination techniques from two devices. The participants controlled the cursor to fly through target rings in the air by the seven travel techniques. In the results, the completion time and success rates of head+eye were very close to head-only, the participants also liked the head+eye technique. But the calibration issue and learning effects noticeably influenced the eye-only input technique, which also

yielded high cyber sickness. This also confirmed that the combination of head and eye worked better and compensated the imprecision of the eye-tracker. In the subjective questionnaires, the participants rated the head-only higher than head+eye, while joystick-based were the worst in travel performance; the participants rated head+eye higher than head-only in NASA-TLX; joystick-based techniques caused the highest cycbersickness, but eye-only caused high sickness symptoms as well, the head+eye was rated closely to head-only.

5.2   Limitations

In the selection study, we necessarily constrained our test conditions to conform to the Fitts' law paradigm. This included keeping participants seated – although we note that seated VR is still a major use case, e.g., gaming on the Oculus Rift. We also constrained targets to only appear in front of the viewer, which is somewhat unrealistic. We considered having targets outside the field of view, but this would not support modeling according to Fitts' law, it would incorporate an extra search task to selection.

In the travel study, we constrained the user's movements always in the air. We considered to implement this testbed because head-based movement supported 6DOF movement of rotation and translation thus we could leverage its advantaged when adding it to 2DOF single techniques (eye, mouse, joystick). However, the single techniques with 2DOF caused certain levels of cycbersickness, especially for eye-only and joystick only. The travel tasks on a surface might be much more comfortable and less nausea for these single techniques, but we did not include it because we wanted to test the most difficult tasks and reveal the most significant results.

In both of our studies, we strongly felt the hardware, software and accessories influenced the eye's performance. Not only the inaccuracy of calibrations, but the weight and belt of the HMD also mattered the process. When we considered the testing equipment, FOVE was the only choice because it launched early when other new products were still in developing. We believe that other HMDs like Tobii in Oculus would have better hardware configuration and calibration mechanism.

5.3   Future Work

In both of our studies, four participants did not pass the calibration or withdrew from the sessions even though they tried many times. Three of them had normal visions, another wore glasses. It is important to mention that two participants who performed pretty well on eye-only and head+only in the studies had rich dancing experience since their childhood. However, we cannot conclude that certain anthropology or ethnography differences caused this issue since we did not have a big sample of participants and all the other participants passed the calibration easily. We cannot make any connections to user's habits since we did not gather enough data before the studies. We hope more studies with larger samples would reveal it in the future.

In the travel study, we observed different levels of learning effect at the end of the sessions on some participants. Because of time and cost limitation, we controlled our experiments in 70 minutes. Some participants took much longer time than others to learn to control their eyes properly. They performed much worse at the beginning than at the end of the sessions, even though they practiced several pre-test trials and the first trials were easier with lower angle deviations. The users need to much more time than we anticipated to learn this new interaction technique. The future work could investigate how

long and how difficult the users would adapt to eye-based interaction technique in VR, what kind of users would need more effort.

The testing results of travel study of eye-based techniques were relatively better than selection because the travel task was more natural based on eye's smooth pursuit. We should conduct researches that follow eye's natural movement mechanism, leverage eye-tracking's advantages[78]. For example, we look at targets at different distances by using eyes' converging and diverging. Then we could develop algorithms to control the cursor in different depth by our eyes.

Future work would also focus on eye-based interaction in VR using a broader range of tasks (e.g., manipulation) and enhanced task realism (e.g., selecting targets outside the field of view).

5.4   Design Recommendations

Based on both of our studies, we propose the following guidelines for designing eye-based VR techniques:

1. Due to eye's microsaccades and FOVE's tracking precision, we recommend target's angular size in FOVE HMD should be greater than 3° to obtain stable performance. Other HMDs with better tracking precision could design greater than 2° of target's angular size. However, we recommend that 3° is the threshold value of adjusting eye ray's sensitives in software development in order to obtain a more stable cursor.

2. We recommend that eye-based technique should combine with other techniques as an input. The head seems a good candidate to collaborate with

eye-tracking when conducting selection and travel based on our results. We do

not recommend single eye input in VR.

Appendices

Appendix A Consent Form

A.1   Consent Form for Selection Study

Carleton
UNIVERSITY
Canada's Capital University

··········**Consent Form**

Title: Evaluation of Selection Performance of FOVE HMD

Funding Source: NSERC

Date of ethics clearance: 01/07/2016

Ethics Clearance for the Collection of Data Expires: 31/12/2018

Researcher: Heather Qian, Carleton School of Computer Science

I _____, choose to participate in a study on evaluation of selection Performance of FOVE HMD. The study will involve participants responding to target spheres in virtual reality by pressing keyboard. **The researcher for this study is Heather Qian, a Masters student in the School of Computer Science.** She is working under the supervision of Dr. Rob Teather in the School of Information Technology at Carleton University.

This experiment involves a 45 mins to one-hour test on computer. You will be wearing FOVE Head-Mounted Display and selecting target objects by eyes and pressing keyboard as the object appears. You should select the object as fast as you can. Software will automatically and anonymously log your task performance data during participation. We may also (with your consent) take video recordings of the screens for further analysis.

Like most virtual reality systems, ours uses a stereo 3D display. Hence you must have normal or corrected-to-normal stereo viewing capabilities to participate—if you can see the 3D effect in 3D movies, you should be fine. There is thus also a minor risk of eye fatigue, headache, or nausea, consistent with the use of stereo displays (e.g., 3D movies) for about 10% of the population. Should you experience such symptoms, you are encouraged to take a break to alleviate the discomfort, or alternatively to withdraw from the study.

You have the right to end your participation in the study at any time, for any reason, up until completion of the experiment. If you withdraw from the study, all information you have provided will be immediately destroyed.

As a token of appreciation, you will receive $10 upon completion of the experiment. This is yours to keep, even if you withdraw from the study.

All research data will be encrypted and stored on the researcher's password-protected hard drive. Any

70

hard·copies·of·data·(including·any·handwritten·notes·or·USB·keys)·will·be·kept·in·a·locked·cabinet·at·Carleton·University.·Research·data·will·only·be·accessible·by·the·researcher.¶

¶

Once·the·project·is·completed,·all·research·data·will·be·kept·in·anonymous·form·indefinitely,·and·potentially·used·for·other·research·projects·on·this·same·topic.·If·you·wish·to·withdraw·your·data·or·any·video·recordings,·please·contact·us·and·we·will·arrange·to·meet·with·you·to·do·so.·If·you·would·like·a·copy·of·the·finished·research·project,·you·are·invited·to·contact·the·researcher·to·request·an·electronic·copy·which·will·be·provided·to·you.¶

¶

The·ethics·protocol·for·this·project·was·reviewed·by·the·Carleton·University·Research·Ethics·Board,·which·provided·clearance·to·carry·out·the·research.·Should·you·have·questions·or·concerns·related·to·your·involvement·in·this·research,·please·contact:¶

¶

CUREB·contact·information:¶
Professor·Shelley·Brown,·Chair·(CUREB-B)¶
Professor·Andy·Adler,·Vice-Chair·¶
Carleton·University·Research·Ethics·Board¶
Carleton·University¶
511·Tory¶
1125·Colonel·By·Drive¶
Ottawa,·ON·K1S·5B6¶
Tel:·613-520-2517¶
ethics@carleton.ca¶

Researcher·contact·information:	**Supervisor·contact·information:**¶
Heather·Qian	Rob·Teather¶
School·of·Computer·Science	School·of·Information·Technology·¶
Carleton·University·	Carleton·University¶

Email:·heatherqian@cmail.carleton.ca··········Email:·RobTeather@cunet.carleton.ca¶

¶
¶

Do·you·agree·to·be·video-recorded·of·the·screen:···___Yes·····___No¶

¶
¶

_____·······_____·¶
Signature·of·participant·······Date¶

¶

_____·······_____·¶
Signature·of·researcher·······Date¶

¶
¶
¶
¶
¶

71

A.2   Consent Form for Travel Study

Carleton UNIVERSITY
Canada's Capital University

----------Consent·Form¶

Title: An·Empirical·Evaluation·of·Eye-Based·Travel·in·Virtual·Reality¶

Funding·Source: NSERC¶

Date·of·ethics·clearance: 01/07/2016¶

Ethics·Clearance·for·the·Collection·of·Data·Expires: 31/12/2018¶

Researcher: Heather·Qian,·Carleton·School·of·Computer·Science¶

I·_____,·choose·to·participate·in·a·study·on·evaluation·the·navigation·performance·of·FOVE·HMD·and·other·input·techniques.··The·study·will·involve·participants·flying·through·rings·in·virtual·reality·by·controlling·an·input·device.·**The·researcher·for·this·study·is·Heather·Qian,·a·Masters·student·in·the·School·of·Computer·Science.**·She·is·working·under·the·supervision·of·Dr.·Rob·Teather·in·the·School·of·Information·Technology·at·Carleton·University.¶

This·experiment·involves·a·60·to·80·min·test·on·laptop.·You·should·fly·through·the·center·of·the·ring·as·closely·as·you·can.··Software·will·automatically·and·anonymously·log·your·task·performance·data·during·participation.·We·may·also·(with·your·consent)·take·video·recordings·of·the·screens·for·further·analysis.¶

Like·most·virtual·reality·systems,·ours·uses·a·stereo·3D·display.··Hence·you·must·have·normal·or·corrected-to-normal·stereo·viewing·capabilities·to·participate—if·you·can·see·the·3D·effect·in·3D·

movies,·you·should·be·fine.·There·is·thus·also·a·minor·risk·of·eye·fatigue,·headache,·or·nausea,·consistent·with·the·use·of·stereo·displays·(e.g.,·3D·movies)·for·about·10%·of·the·population.·Should·you·experience·such·symptoms,·you·are·encouraged·to·take·a·break·to·alleviate·the·discomfort,·or·alternatively·to·withdraw·from·the·study.¶

You·have·the·right·to·end·your·participation·in·the·study·at·any·time,·for·any·reason,·up·until·completion·of·the·experiment.··If·you·withdraw·from·the·study,·all·information·you·have·provided·will·be·immediately·destroyed.¶

As·a·token·of·appreciation,·you·will·receive·$10·upon·completion·of·the·experiment.·This·is·yours·to·keep,·even·if·you·withdraw·from·the·study.¶

All·research·data·will·be·encrypted·and·stored·on·the·researcher's·password-protected·hard·drive.·Any·hard·copies·of·data·(including·any·handwritten·notes·or·USB·keys)·will·be·kept·in·a·locked·cabinet·at·Carleton·University.·Research·data·will·only·be·accessible·by·the·researcher.¶

This·document·has·been·printed·on·both·sides·of·a·single·sheet·of·paper.¶
Please·retain·a·copy·of·this·document·for·your·records.¶

Once the project is completed, all research data will be kept in anonymous form indefinitely, and potentially used for other research projects on this same topic. If you wish to withdraw your data or any video recordings, please contact us and we will arrange to meet with you to do so. If you would like a copy of the finished research project, you are invited to contact the researcher to request an electronic copy which will be provided to you.

The ethics protocol for this project was reviewed by the Carleton University Research Ethics Board, which provided clearance to carry out the research. Should you have questions or concerns related to your involvement in this research, please contact:

CUREB contact information:
Professor Shelley Brown, Chair (CUREB-B)
Professor Andy Adler, Vice-Chair
Carleton University Research Ethics Board
Carleton University
511 Tory
1125 Colonel By Drive
Ottawa, ON K1S 5B6
Tel: 613-520-2517
ethics@carleton.ca

Researcher contact information:	**Supervisor contact information:**
Heather Qian	Rob Teather
School of Computer Science	School of Information Technology
Carleton University	Carleton University
Email: heatherqian@cmail.carleton.ca	Email: RobTeather@cunet.carleton.ca

Do you agree to be video recorded of the screen: ___Yes   ___No

_____   _____
Signature of participant   Date

_____   _____
Signature of researcher   Date

Appendix B Demographic Questionnaires

B.1   Demographic Questionnaire for Selection Study

1. **Age: (Select only one.)** *
   Mark only one oval.

   ◯   18 -29

   ◯   30-40

   ◯   40-60

   ◯   61 or above

   ◯   Prefer not to answer

2. **Gender:** *
   Mark only one oval.

   ◯   Male

   ◯   Female

   ◯   Other: _____

3. **Do you have visual impairments:** *
   Mark only one oval.

   ◯   No

   ◯   Other: _____

4. **What is your domain eye:** *
   Mark only one oval.

   ◯   Left

   ◯   Right

5. **Level of education:** *

Mark only one oval.

- ◯ Some high school but no diploma
- ◯ High school
- ◯ Some college but no diploma
- ◯ College diploma
- ◯ Some university but no degree
- ◯ Bachelor Degree
- ◯ Master's Degree
- ◯ PhD
- ◯ Other: _____

6. **What is your current occupation:** *

_____

7. **Your experience of virtual reality and VR HMD:** *

Mark only one oval.

- ◯ Never
- ◯ Occasional / Less than 5 times in the past
- ◯ Regularly / Few times in a month/year
- ◯ Frequently / More than 10 times in a month
- ◯ Other: _____

B.2   Demographic Questionnaire for Travel Study

1. **Age: (Select only one.)**
   Mark only one oval.

   ◯   18 -29

   ◯   30-40

   ◯   40-60

   ◯   61 or above

   ◯   Prefer not to answer

2. **Gender:**
   Mark only one oval.

   ◯   Male

   ◯   Female

   ◯   Other: _____

3. **What is your current occupation:**

   _____

4. **How often do you use the computer? (desktop, laptop)**
   Mark only one oval.

   ◯   Never

   ◯   Once a month

   ◯   Once a week

   ◯   Daily

   ◯   Other: _____

5. **Your experience of virtual reality and VR HMD:**

   Mark only one oval.

   ◯ Never

   ◯ Occasional / Less than 5 times in the past

   ◯ Regularly / Few times in a month/year

   ◯ Frequently / More than 10 times in a month

   ◯ Daily

   ◯ Other: _____

6. **Which one do you like to play?**

   Tick all that apply.

   ☐ Computer games

   ☐ Mobile games (mobile phone, tablet, watch, etc.)

   ☐ Video games

   ☐ All of them

   ☐ None of them

7. **Your video gaming experience (Xbox, PS4, Nintendo Switch, Wii, 3DS, etc.) :**

   Mark only one oval.

   ◯ Never

   ◯ Occasional / Less than 5 times in the past

   ◯ Regularly / Few times in a month/year

   ◯ Frequently / More than 10 times in a month

   ◯ Daily

   ◯ Other: _____

8. **Do you think you are easy to feel sick/nausea/vomiting?**

Mark only one oval.

- ( ) Never
- ( ) Occasional / when playing in the carnival like riding a roller coaster
- ( ) Sometimes / when riding a vehicle(bus, plane, ship) for a long time
- ( ) Regularly / when riding a vehicle(bus, plane, ship)
- ( ) Other: _____

9. **Do you have visual impairments:**

Mark only one oval.

- ( ) No
- ( ) Yes, I wear glasses.
- ( ) Yes, I wear contact lenses.
- ( ) Yes, I wore glasses/ contact lenses before.
- ( ) Other: _____

Appendix C In-Test Questionnaires

C.1   Device Assessment Questionnaire for Selection Study (after each session of input techniques)

1. Force required for actuation *
Mark only one oval.

	1	2	3	4	5	6	7	
Very uncomfortable	◯	◯	◯	◯	◯	◯	◯	Very comfortable

2. Smoothness during operation *
Mark only one oval.

	1	2	3	4	5	6	7	
Very rough	◯	◯	◯	◯	◯	◯	◯	Very smooth

3. Effort required for operation *
Mark only one oval.

	1	2	3	4	5	6	7	
Very high	◯	◯	◯	◯	◯	◯	◯	Very low

4 Accuracy *
Mark only one oval.

	1	2	3	4	5	6	7	
Very inaccurate	◯	◯	◯	◯	◯	◯	◯	Very accurate

5. Operation speed *
Mark only one oval.

	1	2	3	4	5	6	7	
Unacceptable	◯	◯	◯	◯	◯	◯	◯	Acceptable

6. General comfort *
Mark only one oval.

	1	2	3	4	5	6	7	
Very uncomfortable	◯	◯	◯	◯	◯	◯	◯	Very comfortable

7. Overall operation of input device *

Mark only one oval.

	1	2	3	4	5	6	7	
Very difficult	◯	◯	◯	◯	◯	◯	◯	Very easy

8. Neck fatigue *

Mark only one oval.

	1	2	3	4	5	6	7	
Very high	◯	◯	◯	◯	◯	◯	◯	Very low

9. Eye fatigue *

Mark only one oval.

	1	2	3	4	5	6	7	
Very high	◯	◯	◯	◯	◯	◯	◯	Very low

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

C.3   SSQ for Travel Study (after each session of input techniques)

No_____   Date_____

SIMULATOR SICKNESS QUESTIONNAIRE
Kennedy, Lane, Berbaum, & Lilienthal (1993)***

Instructions : Circle how much each symptom below is affecting you <u>right now</u>.

		None	Slight	Moderate	Severe
1.	General discomfort	None	Slight	Moderate	Severe
2.	Fatigue	None	Slight	Moderate	Severe
3.	Headache	None	Slight	Moderate	Severe
4.	Eye strain	None	Slight	Moderate	Severe
5.	Difficulty focusing	None	Slight	Moderate	Severe
6.	Salivation increasing	None	Slight	Moderate	Severe
7.	Sweating	None	Slight	Moderate	Severe
8.	Nausea	None	Slight	Moderate	Severe
9.	Difficulty concentrating	None	Slight	Moderate	Severe
10.	« Fullness of the Head »	None	Slight	Moderate	Severe
11.	Blurred vision	None	Slight	Moderate	Severe
12.	Dizziness with eyes open	None	Slight	Moderate	Severe
13.	Dizziness with eyes closed	None	Slight	Moderate	Severe
14.	*Vertigo	None	Slight	Moderate	Severe
15.	**Stomach awareness	None	Slight	Moderate	Severe
16.	Burping	None	Slight	Moderate	Severe

* Vertigo is experienced as loss of orientation with respect to vertical upright.

** Stomach awareness is usually used to indicate a feeling of discomfort which is just short of nausea.

C.4   Traveling Performance Questionnaire for Travel Study (after each session of input techniques)

1. **Speed (Actions performable at adequate speed)**
   Mark only one oval.

	1	2	3	4	5	
Do not agree at all	○	○	○	○	○	Completely agree

2. **Accuracy (Actions could be performed precisely)**
   Mark only one oval.

	1	2	3	4	5	
Do not agree at all	○	○	○	○	○	Completely agree

3. **Spatial Awareness (No disorientation after performing a movement)**
   Mark only one oval.

	1	2	3	4	5	
Do not agree at all	○	○	○	○	○	Completely agree

4. **Ease of learning (Interaction was easy to get used to)**
   Mark only one oval.

	1	2	3	4	5	
Do not agree at all	○	○	○	○	○	Completely agree

5. **Ease of use (Actions could be easily performed)**
   Mark only one oval.

	1	2	3	4	5	
Do not agree at all	○	○	○	○	○	Completely agree

Appendix D Post-Test Interviews and Questionnaires

D.1   Interview Questions for Selection Study

1. Overall, was there anything that made you feel uncomfortable during the test? Explain.

_____

_____

2. Which was your favorite selection technique? Why?

_____

_____

3. Which selection technique did you dislike the most? Why?

_____

_____

4. What applications/scenarios would you like to use eye-only technique? Why?

_____

_____

5. What applications/scenarios would you like to use head+eye technique? Why?

_____

_____

D.2   Interview Questions for Travel Study

1. Overall, was there anything that made you feel uncomfortable during the test? Explain.

_____

_____

2. Which was your favorite selection technique? Why?

_____

_____

3. Which selection technique did you dislike the most? Why?

_____

_____

4. What applications/scenarios would you like to use eye-only technique? Why?

_____

_____

5. What applications/scenarios would you like to use head+eye technique? Why?

_____

_____

Overall Satisfaction

Rate your preference

1. **I am very satisfied with the way "mouse" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

2. **I am very satisfied with the way "head" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

3. **I am very satisfied with the way "eye" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

4. **I am very satisfied with the way "joystick" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

5. **I am very satisfied with the way "head+mouse" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

6. **I am very satisfied with the way "head+eye" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

7. **I am very satisfied with the way "head+joystick" performed on this project**
   Mark only one oval.

	1	2	3	4	5	
Strongly Disagree	◯	◯	◯	◯	◯	Strongly Agree

Appendix E Poster Call for Participants

E.1   Posters for Selection Study

Participate in a study on

evaluation of selection performance of FOVE HMD

To participate in this study, you must be:

- ✓ Having normal or corrected vision, and having the ability to see in stereo 3D.
- ✓ Should be adults and English-speaking.

This is a 60-minute session on campus. You will be wearing FOVE Head-Mounted Display and selecting target objects by eyes and pressing keyboard as the object appears. As a token of appreciation, you will receive $10 cash.

The ethics protocol for this project has been reviewed and cleared by the CUREB-B, Protocol #106142. If

you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University

Research Ethics Board-B (by phone at 613-520-2600 ext. 2517 or via email at ethics@carleton.ca).

If you're interested, please contact the researcher, heatherqian@cmail.carleton.ca, for more details.

A VR first-person flying study

To participate in this study, you must be:

- ✓ Having a normal or corrected vision, and having the ability to see in stereo 3D.
- ✓ Gaming experience.
- ✓ Should be adults and English-speaking.

This is a 60 to 80-minute session on campus. FOVE Head-Mounted Display is the world's first commercial eye-tracking VR HMD. You will be wearing FOVE and flying through rings by eyes and other input techniques like a traditional mouse and an Xbox joystick.

As a token of appreciation, you will receive $10 cash.

The ethics protocol for this project has been reviewed and cleared by the CUREB-B, Protocol #106142. If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B (by phone at 613-520-2600 ext. 2517 or via email at ethics@carleton.ca).

If you're interested, please contact the researcher, heatherqian@cmail.carleton.ca, for more details.

References

[1]   P. Fuchs, G. Moreau, and P. Guitton, *Virtual Reality: Concepts and Technologies*. 2011.

[2]   N. Bates-Brkljac, "Virtual reality," *Virtual Real.*, pp. 1–192, 2012.

[3]   G. Burdea and P. Coiffet, "Virtual reality technology," *Presence Teleoperators virtual Environ.*, vol. 12, no. 6, pp. 663–664, 2003.

[4]   S. Currell, "Streamlining the eye : speed reading and the revolution of words, 1870-1940," in *Residual Media*, 2007.

[5]   R. J. K. Jacob and K. S. Karn, "Eye Tracking in Human-Computer Interaction and Usability Research. Ready to Deliver the Promises.," *Mind's Eye Cogn. Appl. Asp. Eye Mov. Res.*, pp. 531–553, 2003.

[6]   R. J. K. Jacob, "What you look at is what you get: eye movement-based interaction techniques," in *Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people - CHI '90*, 1990, pp. 11–18.

[7]   I. S. Mackenzie, "Evaluating Eye Tracking Systems for Computer Input," in *Gaze Interaction and Applications of Eye Tracking: Advances in Assistive Technologies*, 2012, pp. 205–225.

[8]   J. Hales, D. Rozado, and D. Mardanbegi, "Interacting with Objects in the Environment by Gaze and Hand Gestures," *Proc. 3rd Int. Work. Pervasive Eye Track. Mob. Eye-Based Interact.*, pp. 1–9, 2013.

[9]   I. Chatterjee, R. Xiao, and C. Harrison, "Gaze + Gesture : Expressive , Precise and Targeted Free-Space Interactions," *Proc. 2015 ACM Int. Conf. Multimodal Interact.*, no. C, pp. 131–138, 2015.

[10] S. Stellmach, S. Stober, A. Nürnberger, and R. Dachselt, "Designing Gaze-supported Multimodal Interactions for the Exploration of Large Image Collections."

[11] S. Stellmach and R. Dachselt, "Investigating gaze-supported multimodal pan and zoom," *Proc. Symp. Eye Track. Res. Appl.*, vol. 1, pp. 357–360, 2012.

[12] S. Stellmach and R. Dachselt, "Still Looking : Investigating Seamless Gaze-supported Selection , Positioning , and Manipulation of Distant Targets," pp. 285–294, 2013.

[13] A. Patney *et al.*, "Perceptually-based foveated virtual reality," in *ACM SIGGRAPH 2016 Emerging Technologies on - SIGGRAPH '16*, 2016, pp. 1–2.

[14] Y. S. Pai, B. Tag, B. Outram, N. Vontin, K. Sugiura, and K. Kunze, "GazeSim: Simulating Foveated Rendering Using Depth in Eye Gaze for VR," *ACM SIGGRAPH 2016 Posters - SIGGRAPH '16*, pp. 1–2, 2016.

[15] D. Mardanbegi, D. W. Hansen, and T. Pederson, "Eye - based head gestures movements," vol. 1, no. 212, pp. 139–146, 2012.

[16] I. Starker and R. A. Bolt, "A gaze-responsive self-disclosing display," *Proc. SIGCHI Conf. Hum. factors Comput. Syst. Empower. people - CHI '90*, no. April, pp. 3–10, 1990.

[17] K. Essig, D. Dornbusch, D. Prinzhorn, H. Ritter, J. Maycock, and T. Schack, "Automatic analysis of 3D gaze coordinates on scene objects using data from eye-tracking and motion-capture systems," *Proc. Symp. Eye Track. Res. Appl. - ETRA '12*, no. May, p. 37, 2012.

[18] A. T. Duchowski, E. Medlin, A. Gramopadhye, B. Melloy, and S. Nair, "Binocular

eye tracking in VR for visual inspection training," in *Proceedings of the ACM symposium on Virtual reality software and technology - VRST '01*, 2001, p. 1.

[19] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI," *Int. J. Hum. Comput. Stud.*, vol. 61, no. 6, pp. 751–789, 2004.

[20] J. D. Foley, V. L. Wallace, and P. Chan, "HUMAN FACTORS OF COMPUTER GRAPHICS INTERACTION TECHNIQUES.," *IEEE Comput. Graph. Appl.*, vol. 4, no. 11, pp. 13–48, 1984.

[21] M. Mine, "Virtual environment interaction techniques," *UNC Chapel Hill Comput. Sci. Tech. Rep. TR95-018*, pp. 507242–507248, 1995.

[22] I. Poupyrev, S. Weghorst, M. Billinghurst, and T. Ichikawa, "A framework and testbed for studying manipulation techniques for immersive VR," *Proc. ACM Virtual Real. Syst. Technol. Conf. - VRST'97*, pp. 21–28, 1997.

[23] R. J. Teather and W. Stuerzlinger, "Pointing at 3D Target Projections with One-Eyed and Stereo Cursors," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. - CHI '13*, pp. 159–168, 2013.

[24] I. Poupyrev, M. Billinghurst, S. Weghorst, and T. Ichikawa, "The go-go interaction technique," *Proc. 9th Annu. ACM Symp. User interface Softw. Technol. - UIST '96*, pp. 79–80, 1996.

[25] J. S. Pierce, A. S. Forsberg, M. J. Conway, S. Hong, R. C. Zeleznik, and M. R. Mine, "Image plane interaction techniques in 3D immersive environments," *Proc. 1997 Symp. Interact. 3D Graph. - SI3D '97*, p. 39-, 1997.

[26] D. A. Bowman, D. Koller, and L. F. Hodges, "Travel in immersive virtual

environments: an evaluation of viewpoint motion control techniques," in *Proceedings of IEEE 1997 Annual International Symposium on Virtual Reality*, 1997, p. 45–52,.

[27] S. Stellmach and R. Dachselt, "Designing gaze-based user interfaces for steering in virtual environments," *Proc. of ETRA'12*, vol. 1, no. 212, pp. 131–138, 2012.

[28] R. C. Zeleznik, A. S. Forsberg, and J. P. Schulze, "Look-That-There: Exploiting Gaze in Virtual Reality Interactions," *Tech. Rep. CS-05-04*, no. March, 2005.

[29] C. Zanbaka, B. Lok, S. Babu, D. Xiao, A. Ulinski, and L. F. Hodges, "Effects of travel technique on cognition in virtual environments," in *Proceedings - Virtual Reality Annual International Symposium*, 2004, pp. 149–156.

[30] R. Kopper, D. A. Bowman, M. G. Silva, and R. P. McMahan, "A human motor behavior model for distal pointing tasks," *Int. J. Hum. Comput. Stud.*, vol. 68, no. 10, pp. 603–615, 2010.

[31] M. F. Stoelen and D. L. Akin, "Assessment of fitts law for quantifying combined rotational and translational movements," *Hum. Factors*, vol. 52, no. 1, pp. 63–77, 2010.

[32] Y. Qian and R. J. Teather, "The Eyes Don ' t Have It : An Empirical Comparison of Head -Based and Eye-Based Selection in Virtual Reality," pp. 1–8, 2017.

[33] R. L. Gregory, "Eye and brain : the psychology of seeing," *Br. Med. J.*, vol. 4, no. 5893, p. 682, 1973.

[34] G. Alder, "Psychology: The Science of Behaviour.," *Can Psychol*, vol. 43, no. 4, pp. 278–280, 2002.

[35] J. M. Findlay, "Saccadic eye movement programming: Sensory and attentional

factors," *Psychol. Res.*, vol. 73, no. 2, pp. 127–135, 2009.

[36]   G. Westheimer and S. P. McKee, "Visual acuity in the presence of retinal-image motion," *J. Opt. Soc. Am.*, vol. 65, no. 7, pp. 847–850, 1975.

[37]   J. Lanman, E. Bizzi, and J. Allum, "The coordination of eye and head movement during smooth pursuit," *Brain Res.*, vol. 153, no. 1, pp. 39–53, 1978.

[38]   R. A. Bolt, "Gaze-orchestrated dynamic windows," *ACM SIGGRAPH Comput. Graph.*, vol. 15, no. 3, pp. 109–119, 1981.

[39]   R. A. Bolt, "Eyes at the interface," *Proceedings of the 1982 conference on Human factors in computing systems*. p. , 1982.

[40]   S. Stellmach and R. Dachselt, "Look & Touch : Gaze-supported Target Acquisition," 2012.

[41]   J. Turner, J. Alexander, A. Bulling, and H. Gellersen, "Gaze+RST: Integrating Gaze and Multitouch for Remote Rotate-Scale-Translate Tasks," *Proc. ACM CHI'15 Conf. Hum. Factors Comput. Syst.*, vol. 1, pp. 4179–4188, 2015.

[42]   K. Pfeuffer and H. Gellersen, "Gaze and Touch Interaction on Tablets," *Proc. 29th Annu. Symp. User Interface Softw. Technol. - UIST '16*, pp. 301–311, 2016.

[43]   K. Pfeuffer, J. Alexander, and H. Gellersen, "Partially-indirect Bimanual Input with Gaze, Pen, and Touch for Pan, Zoom, and Ink Interaction," *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst. - CHI '16*, pp. 2845–2856, 2016.

[44]   G. Fabian, S. Vogt, S. Stellmach, and A. Siegel, "Gaze-supported Foot Interaction in Zoomable Information Spaces," 2013.

[45]   A. T. Duchowski, V. Shivashankaraiah, T. Rawls, A. K. Gramopadhye, B. J. Melloy, and B. Kanki, "Binocular eye tracking in virtual reality for inspection

training," *Proc. 2000 Symp. Eye Track. Res. Appl.*, pp. 89–96, 2000.

[46]   L. Vanacken, T. Grossman, and K. Coninx, "Exploring the effects of environment density and target visibility on object selection in 3D virtual environments," in *IEEE Symposium on 3D User Interfaces 2007 - Proceedings, 3DUI 2007*, 2007, pp. 115–122.

[47]   P. Isokoski and B. Martin, "Eye Tracker Input in First Person Shooter Games," pp. 4–6, 2006.

[48]   J. D. Smith and T. C. N. Graham, "Use of Eye Movements for Video Game Control," 2006.

[49]   T. Ohshima, H. Yamamoto, and H. Tamura, "Gaze-directed adaptive rendering for interacting with virtual space," in *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, 1996, pp. 103–110.

[50]   W. Steptoe *et al.*, "Eye-tracking for avatar eye-gaze and interactional analysis in immersive collaborative virtual environments," *Proc. ACM 2008 Conf. Comput. Support. Coop. Work - CSCW '08*, no. November 2008, p. 197, 2008.

[51]   L. E. Sibert and R. J. K. Jacob, "Evaluation of eye gaze interaction," *Proc. CHI*, vol. 2, no. 1, pp. 281–288, 2000.

[52]   D. Fono and R. Vertegaal, "EyeWindows: evaluation of eye-controlled zooming windows for focus selection," *Proc. SIGCHI Conf. Hum. factors Comput. Syst.*, pp. 151–160, 2005.

[53]   P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, pp. 381–391, 1954.

[54]   I. Standard, "Ergonomic requirements for office work with visual display terminals

(VDTs) - Part 9: Requirements for non -keyboard input devices," *Iso 2000*, vol. 2000, p. 54, 2000.

[55] I. S. MacKenzie, "Fitts' Law as a Research and Design Tool in Human-Computer Interaction," *Human–Computer Interact.*, vol. 7, no. 1, pp. 91–139, 1992.

[56] I. S. MacKenzie and W. Buxton, "EXTENDING FITTS' LAW TO TWO-DIMENSIONAL TASKS INTRODUCTION Since the advent of direct manipulation," *ACM CHI'92 Conf.*, pp. 219–226, 1992.

[57] R. Vertegaal, "A Fitts Law comparison of eye tracking and manual input in the selection of visual targets," *Proc. 10th Int. Conf. Multimodal interfaces - IMCI '08*, p. 241, 2008.

[58] D. Bowman, E. Kruijff, J. LaViola, and I. Poupyrev, *3D user interfaces: theory and practice*. 2004.

[59] S. Lee, J. Seo, G. J. Kim, and C.-M. Park, "<title>Evaluation of pointing techniques for ray casting selection in virtual environments</title>," no. January 2014, pp. 38–44, 2003.

[60] R. J. Teather and W. Stuerzlinger, "Pointing at 3D targets in a stereo head-tracked virtual environment," *3DUI 2011 - IEEE Symp. 3D User Interfaces 2011, Proc.*, no. 1, pp. 87–94, 2011.

[61] E. R. F. W. Crossman and P. J. Goodeve, "Feedback control of hand-movement and Fitts' Law," *Q. J. Exp. Psychol. Sect. A*, vol. 35, no. 2, pp. 251–278, 1983.

[62] V. Tanriverdi and R. J. K. Jacob, "Interacting with eye movements in virtual environments," *Proc. SIGCHI Conf. Hum. factors Comput. Syst. - CHI '00*, vol. 2, no. 1, pp. 265–272, 2000.

[63] N. Cournia, J. D. Smith, and A. T. Duchowski, "Gaze- vs. Hand-based Pointing in Virtual Environments," *Proc. CHI '03 Ext. Abstr. Hum. Factors Comput. Syst. (CHI '03)*, vol. 2, p. 772, 2003.

[64] R. G. Golledge, "Human wayfinding and cognitive maps," in *The Colonization of Unfamiliar Landscapes: The Archaeology of Adaptation*, 2003, pp. 25–43.

[65] W. T. Nelson, L. J. Hettinger, J. A. Cunningham, M. M. Roe, M. W. Haas, and L. B. Dennis, "Navigating through virtual flight environments using brain-body-actuated control," *Proc. IEEE 1997 Annu. Int. Symp. Virtual Real.*, pp. 30–37, 1997.

[66] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness," *Int. J. Aviat. Psychol.*, vol. 3, no. 3, pp. 203–220, 1993.

[67] J. J. LaViola, "A discussion of cybersickness in virtual environments," *ACM SIGCHI Bull.*, vol. 32, no. 1, pp. 47–56, 2000.

[68] E. M. Kolasinski, "Simulator Sickness in Virtual Environments," *United States Army Research Institute fo the Behavioral and Social Sciences*, vol. 1027, no. 4. p. 68, 1995.

[69] D. M. (U. S. A. R. I. Johnson, "Introduction to and Review of Simulator Sickness Research," *Soc. Sci.*, no. April, 2005.

[70] L. J. Hettinger, K. S. Berbaum, R. S. Kennedy, W. P. Dunlap, and M. D. Nolan, "Vection and simulator sickness.," *Mil. Psychol.*, vol. 2, no. 3, pp. 171–181, 1990.

[71] R. H. So, W. T. Lo, and a T. Ho, "Effects of navigation speed on motion sickness caused by an immersive virtual environment.," *Hum. Factors*, vol. 43, no. 3, pp.

452–461, 2001.

[72] W. Chen, A. Plancoulaine, N. Férey, D. Touraine, J. Nelson, and P. Bourdot, "6DoF Navigation in Virtual Worlds: Comparison of Joystick-based and Head-controlled Paradigms," *Proc. 19th ACM Symp. Virtual Real. Softw. Technol. - VRST '13*, p. 111, 2013.

[73] X. Zhang and I. S. MacKenzie, "Evaluating eye tracking with ISO 9241-Part 9," in *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, 2007, pp. 779–788.

[74] I. S. Mackenzie and R. J. Teather, "FittsTilt : The Application of Fitts ' Law To Tilt-based Interaction," in *7th Nordic Conference on Human-Computer Interaction*, 2012, pp. 568–577.

[75] M. Brown, W. Stuerzlinger, and E. J. Mendonça Filho, "The performance of un-instrumented in-air pointing," *Proc. 2014 Graph. Interface Conf.*, pp. 59–66, 2014.

[76] B. Ashworth, "Movements of the Eyes," *J. Neurol. Neurosurg. Psychiatry*, vol. 41, no. 1599, p. 769, 1978.

[77] C. T. Brandt, "Subcortical Control of Visual Fixation," 2011.

[78] W. Robert, M. Exploring, T. Piumsomboon, and R. W. Lindeman, "Exploring natural eye-gaze-based interaction for immersive virtual reality This is the peer reviewed author accepted manuscript (post print) version of a published work that appeared in final form in : This un-copyedited output may not exactly replicate," 2017.