Western University

Scholarship@Western

Undergraduate Honours Theses

Psychology

Spring 5-15-2020

DEPTH PERCEPTION IN VIRTUAL PERIPERSONAL SPACE: AN INVESTIGATION OF MOTION PARALLAX ON PERCEPTION- VS ACTION-ESTIMATIONS

Hongyao Zhu

DEPTH PERCEPTION IN VIRTUAL PERIPERSONAL SPACE:

AN INVESTIGATION OF MOTION PARALLAX ON PERCEPTION- VS ACTION-ESTIMATIONS

by

Hongyao Zhu

Department of Psychology

Submitted in Partial Fulfillment

of the requirements for the degree of

Bachelor of Arts in Honours Psychology

Faculty of Arts and Social Science

Huron University College

London, Canada

April 24, 2020

Abstract

The goal of the current experiment was to investigate whether the addition of Motion Parallax will allow participants to make more accurate distance estimations, in both the real and virtual worlds, as well as to determine whether perception- and action-estimations were affected similarly. Due to rising number of COVID-19 cases in 2020, all in-person testing needed to cease with only one participant being tested with the full set of conditions in the final experimental configuration and one participant having been completed the motion parallax conditions only. As a result, the two participants were combined and only the motion parallax conditions were analyzed. Due to low statistical power, no significant main effects, nor significant interactions were discovered. Once the COVID-19 pandemic has subsidised, I am intending to collect data from all twenty-four participants with the full array of conditions in order to complete the current project. An increase in distance-estimation accuracy, especially in virtual reality conditions is still expected to be found.

Keywords: Virtual Reality, Motion Parallax, Distance Estimation

Acknowledgements

First and foremost, I would like to express my immense gratitude to my thesis advisor, Dr. Derek Quinlan, for taking me on as a thesis student this past year. Your support, guidance, encouragement, and continued feedback on my work are what made this thesis possible. Thank you for always believing in me, giving me the opportunity to explore, and motivating me to learn. Words cannot express how grateful I am to you for your mentorship, and I look forward to learning more from you in the future.

To my second reader, Dr. Christine Tsang. Thank you for your insightful comments as a second reader and being such a compassionate chair of the department. This publication would not have been possible without you.

To Karsten Babin and Kevin Stubbs. Thank you for putting so much work and effort into this project from building software, to writing scripts and data collection. Thank you both for always being so supportive and available for me.

To my other psychology professors, Dr. Dumas and Dr. Cheung, thank you for enriching my undergraduate years and encouraging me to learn.

To Ella and Michael, thank you for being my home far away from home. Thank you for making me feel loved and supported here in Canada.

To all my friends, English teachers and leadership groups at Huron. Thank you for being my inspiration, motivation and making my dreams come true.

Finally, to my parents. Without your support, I would not have been where I am today. Thank you for always trusting me, supporting me, and respecting my every decision. I am so grateful.

Table of Contents

Page

Introduction

In today's technologically advanced world, Virtual Reality (VR) is one of the fastest-growing technologies that will change the future (Oculus, 2019). The idea of VR, or the presentation of stereoscopic visual images, has a long history and, can be traced back to the early days of photography. An example of this historical "VR" is the Stereoscope (Figure 1), which dates back to the early 1830's, which is a binocular apparatus for viewing a pair of stereoscopic images. The stereoscopic images, or photo pairs, are photographed using two cameras set apart laterally by the distance between the average person's pupils (~6 cm).

Figure 1. An image of a stereoscope.

These offset cameras reproduce the respective views of the left and right eyes such that when viewed through a stereoscope, which optically superimposes the two images atop one another, a three-dimensional perception occurs (Figure 2, Parmeggiani & Parmeggiani, 2016). Unfortunately, little research was paid to these stereoscopic images and methods. It was not until over one hundred years later that, the concept of VR was introduced by Sutherland (1965) and the first Head-Mounted Displays (HMD) were invented. One of the unique features of VR, beyond that seen in the earlier stereoscopes, is that it allows individuals to experience visual situations/simulations through the Head-Mounted Displays (HMD), which update their displays based on user's own head movements. VR creates a dynamic scene, different than what you can get from a stereoscopic image, which is static, and giving users the impression of being immersed in an virtual environment (termed as telepresence, Napieralski et al., 2011).

Image 2. An example of stereoscopic image. Adapted from Parmeggiani & Parmeggiani (2016).

Currently, virtual reality has seen an incredible expansion in the entertainment/gaming industry (Oculus, 2019), that many professions, such as designers/engineers, military/police forces, and medical personnel have greatly benefited from by using VR as a training tool (Naceri, Chellali & Hoinville, 2011). For instance, training of cardiac surgeons within VR has allowed them to reach, grasp and manipulate virtual objects, such that they can practice their medical skills within a controlled and safe virtual environment, without placing patients at risk (Peters et al., 2008). As the goal of VR training is to prepare professionals for real world tasks and reduce the probability of errors (Seymour, 2008), it is crucial that VR faithfully recreate all of the cues to depth perception that are present in the real world. (Hoffman, Girshick, Akeley, & Banks, 2008).

While this faithful reproduction has been the goal of VR systems for decades, recent research related to using VR as a training tool has suggested that current hardware is still plagued by depth perception errors. For example, Aggarwal, Black, Hance, Darzi, and Cheshire (2006) investigated the effectiveness of VR simulation training on endovascular skills between eight expert surgeons (> 50 endovascular procedures) versus twelve surgeons with limited experience (< 10 endovascular procedures). All twenty surgeons performed a renal artery balloon angioplasty and stent procedure, with “extensive experienced” surgeons performing two sessions, whilst “inexperienced surgeons” performed six sessions. Overall, VR training was shown to be an effective tool for improving “inexperienced surgeon’s” endovascular skills as measured by a reduction in procedure time (Aggarwal at al., 2006). It should be noted, clinically relevant parameters, such as the accuracy of stent placement and sizing were not measured, two measures that should have been investigated in order to determine the effectiveness of VR surgical training. It was noted that even the “experienced” surgeons needed two sessions to effectively demonstrate their endovascular skills (Aggarwal at al., 2006). Thus, if the VR simulation was able to faithfully replicate the surgical conditions, including depth information cues, then the “experienced” surgeons should not have needed two sessions to adjust to the VR environment.

Similarly, Dayal et al. (2004) looked at the application of VR for training novice versus experienced surgeons in catheter-based skills. They discovered that within a total of twenty-one surgeons (five experienced and sixteen novice), although the time to complete a clinical scenario for novice surgeons had greatly improved after the training program, their time usages (23 minutes) were still greater than those of expert surgeons (13 minutes) (Dayal et al., 2004). Surprisingly, even the experienced surgeons seemingly did not benefit from the VR training program (i.e., shortened time usage), which may be explained by the fact for not having enough clinical and tactile feedback from the VR simulator and the flaws embedded in VR rendering.

Based on those results from the application of VR for medical training, even experienced surgeons seemed to not benefit from the training program (i.e., needed more training sessions to demonstrate their skills). As the current VR training program is still plagued with flaws of VR rendering (i.e., depth perception errors), one might reconsider adopting VR as a training tool for surgical training in an effort to prevent the potentiality of putting patients at unnecessary risk. Therefore, it is crucial for researchers to fully understand the depth perception errors within current VR hardware before it is implemented for extensive training purposes.

In order to investigate the depth perception errors within the current VR surgical training programs, it is important to narrow our focus down to studies that include peripersonal space, the distance within an individual's reach without locomotion (Naceri et al., 2011). Armbürster et al. (2008) investigated depth perception in virtual environments by manipulating the aspects of: 1) the virtual world (no space vs. open space vs. closed space), 2) target distances (varied from 40cm to 500cm), 3) the existence of a metric aid (with vs. without tape measure), and 4) the type of object presentation (single vs. ten). Participants' depth perception accuracy was determined by having the participants verbally reported object distance in centimeters. Although the quality of the virtual world and the existence of a metric aid were shown to have no impact on participant's depth perception, participants did however, have a general tendency to overestimate target distance within peripersonal space, demonstrating a depth perception issue.

Similarly, Murgia and Sharkey (2009) investigated distance estimation in virtual environment by using a virtual-matching task. First, participants were asked to study the size of a real cube and the size of a real sphere in order to establish an idea of the relative dimensions of the object for later testing. Then, participants were asked to stand in the centre of a virtual CAVE (Cave Automatic Virtual Environment), where they viewed a virtual cube and a virtual sphere of the same dimensions. During the experiment, both real and virtual cubes and spheres disappeared, and participants were instructed to use a hand-held joystick to indicate the position where the virtual cube had appeared. In contrary to Armbürster et al. (2008), an overall underestimation of distances within the virtual environment was found amongst participants.

The VR depth perception literature, as limited as it is, seems to be rife with distance estimation errors that are seemingly contradictory. As these studies vary greatly in the methodology used for estimations (verbal, action, joystick-controlled pointers, etc.) and the distances tested, it may be that both of these factors have been contributing to the contradictory findings. In fact, distance does seem to have an effect on estimation error. Naceri et al. (2011) had participants indicate the location of a previously viewed object using their index finger. Interestingly, participant's distance estimation errors were small when objects presented at distances less than 55cm; however, when target's distances exceeded 55 cm, participants made more mistakes and tended to underestimate distances (Naceri et al., 2011). This distance effect could partially explain the contradiction in previous findings; however, addressing the estimation methodologies used previously, should likewise provided a clearer, more accurate "picture" of the depth perception issues plaguing virtual reality HMDs.

It is clear that a depth perception problem persists, but the results of previous research have lacked a general consensus. As mentioned above, this is likely due to the various, artificial means by which distance estimations have been reported. To address the artificiality employed in previous methodologies, it has been suggested that reaching and grasping should be used as the reporting method of distance perceptions (Lockwood, 2017), at least for those distances within peripersonal space. However, two distinct visual pathways have been found to have different functions towards visual stimuli, with the ventral visual pathway processes perceptual information for the purpose of identifying an object. Whereas, the dorsal visual pathway serves the purpose of using visual information to guide bodily movement (Goodale & Milner, 1992). Consequently, due to the existence of two visual pathways, various reporting methods might affect the accuracy of distance estimations in VR (Lockwood, 2017).

To this end, a previous honor thesis project at Huron University College (Lockwood, 2017) investigated depth perception in both the real and virtual worlds by having participants perform what was referred to as “perception”- and “action-based” estimates within peripersonal space. To achieve these ethologically valid estimates, participants either indicated object distance by using the distance between their index fingertips (perception-based estimate) or by reaching out with an index fingertip, as if trying to touch the object (action-based estimate). It was determined that distance estimation errors occurred more so in the virtual world than in the real world, that perception-based estimations were worse than action-based estimations, and that the estimation errors grew as a function of distance from the head, with the worst errors being for objects at arm’s length in VR (Lockwood, 2017). These results demonstrated that even when using the most recently produced HMDs, spatial misrepresentation still exists and should be of concern when being used to train professionals where depth perception is critical (e.g. surgeons).

It is likely that these distance misrepresentations are due to a conflict between the depth-related cues of the Accommodative Reflex. (a.k.a. the Near Triad). This reflex links vergence eye position, pupil size and accommodative state, such that they change in coordination with one another as a person views objects at varying distances. For example, if you were reading your book (which is sitting at arms length), your eyes would be heavily converged (rotated inward), your lenses would be thickened to increase their refractive power, and your pupils would be more constricted. If you were then to look at the moon up in the sky. Your eyes would diverge, rotating until their respective views are parallel, your lenses would thin considerably, reducing their refractive power, and your pupils would dilate. The information from the accommodative and vergence eye position systems serve as depth cues and allow for accurate calculations of object distances (Goldstein, 2013). Due to the fact that all previously and currently manufactured HMDs only manipulate vergence eye position (same as stereoscopes from a hundred years ago) to create a perception of depth, the maintaining of a constant accommodative state, this pits the two major binocular depth cues against one another. This depth cue conflict is easily witnessed when viewing a 3D movie and you notice that you are having some difficulty focusing on objects represented at different distances. As your eyes rotate to verge to another distance, your eyes try to accommodate to that distance as well, just like while reading the book, then looking at the moon. However, the movie screen does not change distance, therefore your visual system must fight the Accommodative Reflex in order to maintain accommodation to only one distance while, despite all of the vergence eye movements. The exact same scenario is present in a HMD, albeit on a smaller scale. If the visual system uses the binocular depth cues of vergence eye position and accommodative state to determine object distance, then holding one of those cues at a constant is going to introduce error into depth perception. To counter this issue, the ideal HMD would not only use vergence eye position to create depth perception, but also alter the HMD optics such that the lenses of viewer’s eyes would need to accommodate appropriately, thereby maintaining the Accommodative Reflex and producing more accurate depth/distance estimates.

It is not only these two binocular cues that inform the visual system as to object distance, there are also numerous monocular cues as well, such as Motion Parallax. Motion Parallax is a depth perception cue that stems from the lateral head movements made by the viewer. More specifically, as a person’s head moves sideways, objects that are closer to the viewer, appear to move faster (side-to-side) than do objects that are further away (Goldstein, 2013). For example, if you were sitting in a moving train and staring out of the side-window, your head would be moving sideways and closer objects such as the telephone posts nearest to the tracks, would appear to move rapidly by, whereas the mountains in the distance move much slower, whilst the sun would be perceived as not moving at all. This amount of side-to-side movement that differs as a function of object distance from the viewer has not been properly investigated in the VR depth perception literature (Kongsilp & Dailey, 2017). Moreover, many previous studies, such as Lockwood's (2017) used head-fixed participants or used static images (i.e. stereoscopic images) in the virtual environment; therefore, eliminating motion parallax.

The goal of current study is to investigate whether the addition of Motion Parallax will allow participants to make more accurate distance estimations, in both the real and virtual worlds, as well as determine whether perception- and action-based estimates are affected similarly. It has been previously determined that even information stemming from microparallax (tiny postural adjustments of only a few millimeters) can be an important depth cue (Tiron & Langer, 2018). Accordingly, the current study will have participants make distance estimates while either being allowed to move their head laterally (i.e. motion parallax cues present), or while being head-fixed (i.e. no motion parallax cues).

Method

1. Participants

1.1) Participant Characteristics

Originally, this experiment required twenty-four participants for proper counterbalancing; however, due to the COVID -19 virus situation of 2020, only two participants took part in the finalized version of this study prior to Huron University and Western University required that “in-person” participant testing be ceased. Therefore, with only two participants’ data collected, the corresponding lack counterbalancing and statistical power dictated that we treat these findings as “pilot” data and will be discussed as such.

1.2) Inclusion/Exclusion Criteria

To ensure the validity of the data and the safety of those participating, several inclusion /exclusion criteria were implemented. First, participants were required to be right-handed to allow them to effectively interact with the testing apparatus (Appendix A). Second, participants were required to have normal or corrected-to-normal visual acuity (contact lenses or corrective surgery only), so that they could clearly see the target objects in all conditions. Those individuals whom only had glasses to correct their vision, they were excluded from the study, as glasses do not fit comfortably within the head mounted displays (HMDs) and may damage the HMD lenses. Third, participants were tested using the RANDOT Stereotests (Stereo Optical Company Inc.) to ensure that they had “normal” stereoacuity (i.e. depth perception; Figure 1). Similar to what has been used by others conducting depth/distance estimation research, such as Fawcett and Birch (2000), only participants with stereoacuity equal to or greater than 40 seconds of arc were allowed to proceed to the testing phase. Fourth, participants with a history of seizures/epilepsy were excluded from the study, as it has been previously reported that flashing images from HMDs may trigger seizures in persons suffering from photosensitive epilepsy (da Silva & Leal, 2017). Lastly, any participant whom wore mascara to the testing session was excluded from the experiment, as mascara is difficult to remove from, and can potentially damage, the lenses of the HMDs.

Figure 1. RANDOT Stereotests (Stereo Optical Company Inc.). Only the ten Circles Stereotests (upper left) were used in the screening process.

1.3) Participant Recruitment

It was initially proposed that participants would be recruited either from the Psychology Research Participation Pool (SONA) at Huron University College or from the family and friends of the researchers. Neither pool of participants had any previous knowledge of the hypotheses. Both male and female participants were qualified for recruitment in the current study. Participants from the Research Pool would receive 1.0 research credit toward their Psychology course (Psych1100E or Psych1000).

2. Research Design

The current experiment manipulated four variables: Environment (Virtual Reality versus Real World), Estimation Method (Perception-Estimation versus Action-Estimation), Motion Parallax (Head Fixed versus Head Non-Fixed) and Target Distance (30 cm, 45 cm, and 60 cm). The combination of those three variables led to eight different testing conditions. The condition orders for each participant were predetermined and designed using a MATLAB script, such that no one condition would follow any other condition systematically in order to control order effects. All estimation tasks (i.e. Action- and Perception-Estimation) were performed open-loop, such that no visual, haptic, nor terminal feedback were available to participants as they made their distance estimations.

The experiment was a within-subject design; therefore, each participant was tested in all eight conditions. Target distances (30cm, 45cm, 60cm) presented within each condition were also completely randomized (i.e. Virtual Reality: Perception-Estimation: Head-Fixed: 30cm, etc.).

3. Materials and Apparatus

3.1) Real-World Stimuli

The Real-World stimuli were three circular “bull’s-eye” targets set at 30, 45, and 60cm from the participants’ eyes (Figure 2). To ensure the accuracy of presentation distances, viewing distance was always measured prior to testing commencement. The diameter of the targets increased with egocentric distance such that the sizes of targets at different physical distances appeared the same on participants’ retina (i.e. retinal equivalence), such that object size could not be used as a cue to distance.

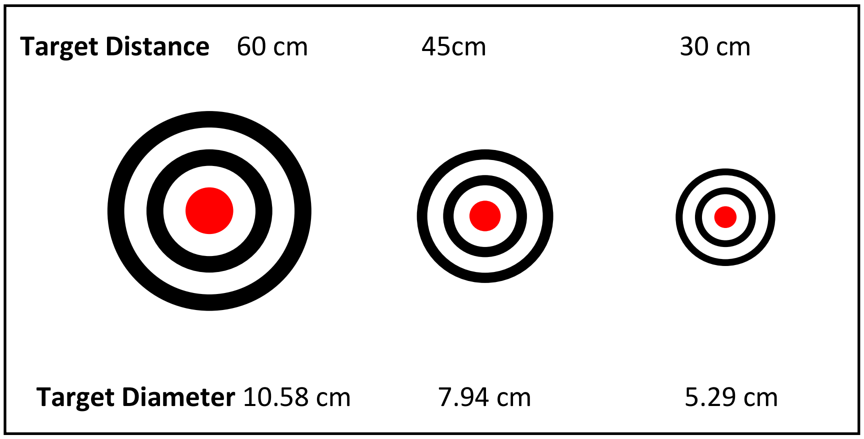


Figure 2. Target egocentric distances and corresponding diameters. Images are not to scale.

Each circular target was glued to a plastic rod and inserted manually into a wooden dowel, which was secured onto a black-coated, wooden testing board (Figure 3). Insertion into dowel allows for swapping targets, dowel pivots such that target can be moved out of the way as participant reaches out during Action Estimation. Therefore, no accuracy feedback from touching the target.

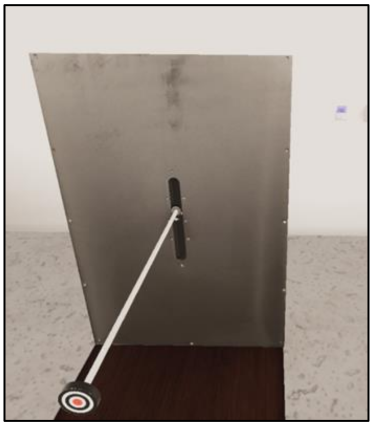


Figure 3. This image shows how a 30 cm distance target looks during testing (from above and to the right of the participant).

3.2) Virtual Stimuli

The virtual testing environment was created using Unity, which is a software tool used for creating video games and other 3D interactive applications/ environments. This program allows users to combine the code and 3D models they have created elsewhere (e.g., Visual Studio and Blender), then combine these together to create a virtual environment.

3.2.1) Virtual Testing Room: The version of Unity used to make the application was Unity 2019.2.6f1. The virtual scene (Figure 4) was constructed by taking individual 3D model objects and aligning them in such a way as to reproduce the real-world testing room and equipment. Some of the objects used to create the scene were custom-made in Blender and Photoshop (e.g., target stimuli, black target backboard, light switch and thermostat). While other objects such as, the chair, table, walls, and floor were purchased from the Unity Asset Store. All virtual items were scaled to precisely match their real-world counterparts in order to faithfully reproduce the testing room, such that rendering size errors could not be used as an explanation to any distance estimation errors.

Figure 4. An image of the virtual scene from the view of Unity editor camera.

3.2.2) Virtual Targets and Other Stimuli: Objects such as the target stimuli, black target board, and thermostat had to be custom made as these items were very specific to our testing room (Figure 5). To create these custom assets for use in building the virtual Unity environment, Photoshop, and Blender were employed.

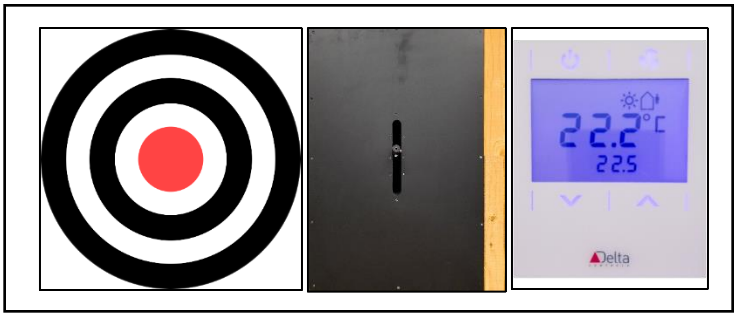


Figure 5. All three custom diffuse texture images were created or altered in Photoshop, then modeled in Blender.

Photoshop (Adobe Inc.) was first used to produce our custom assets, as it was required to either create or edit images that would become the diffuse texture for our 3D models. A diffuse texture is an image that holds all the color information of an object. Once created, the diffuse texture image was exported from Photoshop as a PNG file to be applied to the surface of a 3D model in Blender (Figure 6).

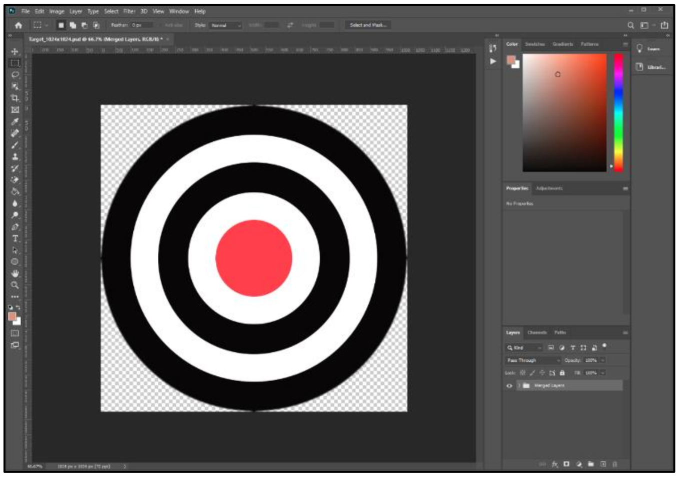
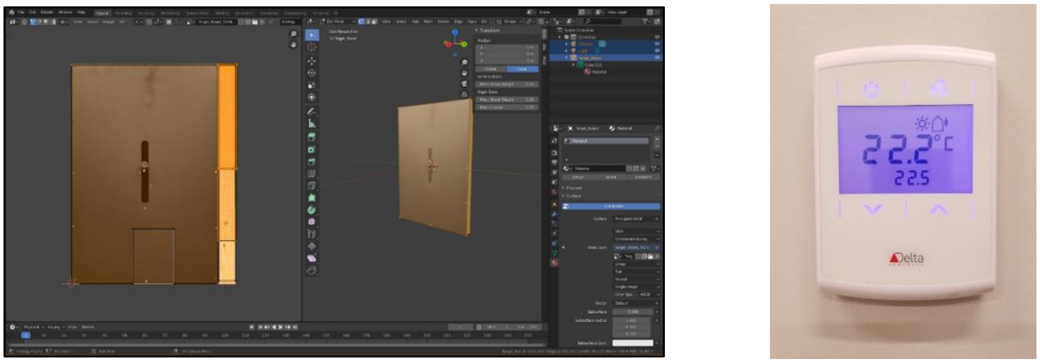


Figure 6. View from within Photoshop of target stimuli diffuse texture being created.

Blender (The Blender Foundation) was the last step in creating our custom 3D assets. In Blender (2.82a), we created simple 3D objects such as cylinders for targets, and a rectangular cube for the black target backboard (Figure 7 a.). The thermostat required a bit more work by needing slopped edges around its primarily square body (Figure 7 b.). Blender also allowed us to set the size of our models based off our real-world measurements to help maintain accuracy.



a. b.

Figure 7.a) An example of creating the black target backboard in Blender, a diffuse texture (i.e., black testing board) was later applied to the rectangular cube above. b) A picture of the thermostat

3.2.3) Dimensions and Size Accuracy: To assure that the scale of the scene was accurately recreated within Unity, measurements were taken of all critical objects and structural elements within the room. All dimensions were recorded in centimeters, as this would provide for the easiest conversion to Unity distance units, which is equal to 1-meter. For example, when setting the scale values for the black target backboard in Unity units, its values would be x = 0.61 y = 0.91 z = 0.06 (Figure 8).

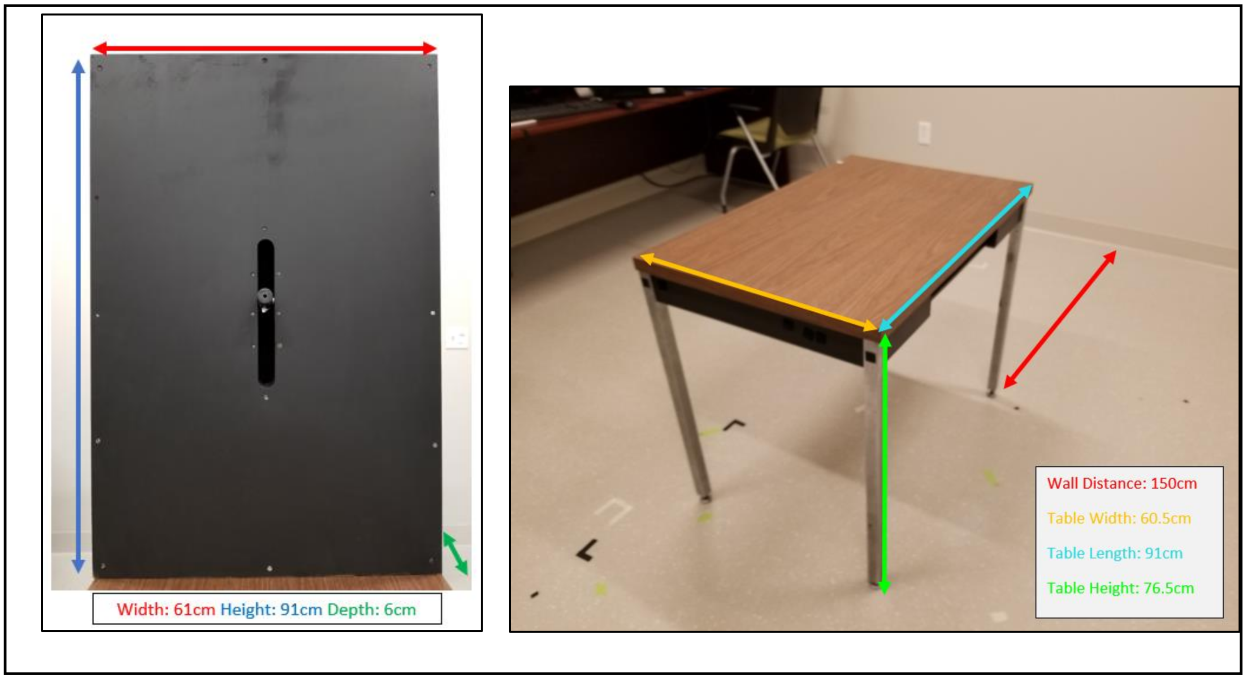
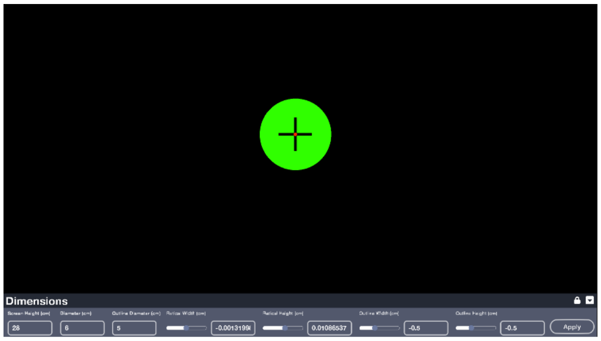


Figure 8. Image on the left demonstrates the black target backboard with real dimensions and image on the right illustrates the table with real dimensions.

3.2.4) After-Image Paradigm: To ensure that virtual objects were rendered to accurate size when viewed in the Oculus Rift CV1 HMD, we created a scale testing in Unity, which we refer to as the After-Image Paradigm. The application creates an after-image in the real world that is then compared against the size of a VR calibration object. If the size of the afterimage was equal to the VR calibration object, then we could be assured that VR scenes were being rendered size accurate. If these images were not of equal size, then scaling factors could be applied until size calibration was achieved.

The application required a flat-screen computer monitor, and the Oculus Rift CV1 HMD be connected to the same computer. The computer monitor rendered a bright green 30mm circular stimulus that would create an afterimage after a period of visual adaptation. Also, the computer monitor is a real-world object, we could verify the actual real-world size of our adaptation stimulus. To create an effective after image, the bright green 30 mm adaptation circle (with fixation cross) was rendered against a black background on the computer monitor (Figure 9).



Screen Height (cm) 28
Diameter (cm) 6
Outline Diameter (cm) 5
Reticle Width (cm) -0.0013199
Reticle Height (cm) 0.01086537
Outline Width (cm) -0.5
Outline Height (cm) -0.5

Figure 9. The interface and adaptation stimuli on the computer monitor.

Within the VR HMD, a 30 mm black ring with a fixation dot at it’s geometric centre, was presented centrally on a white background for viewing (See Figure 10). Viewing this calibration stimulus allowed for the afterimage of the real-world afterimage to be superimposed onto the VR ring. If the afterimage filled the ring exactly, then this procedure verifies that the VR environment was rendering size accurate. However, if the afterimage failed to fill the VR ring stimulus, or extend beyond its’ boundary, then scaling modifications would be necessary within Unity to make the VR environment veridical. This process was only needed to be performed once prior to any testing, after which the system was calibrated, but could be verified throughout the testing process.

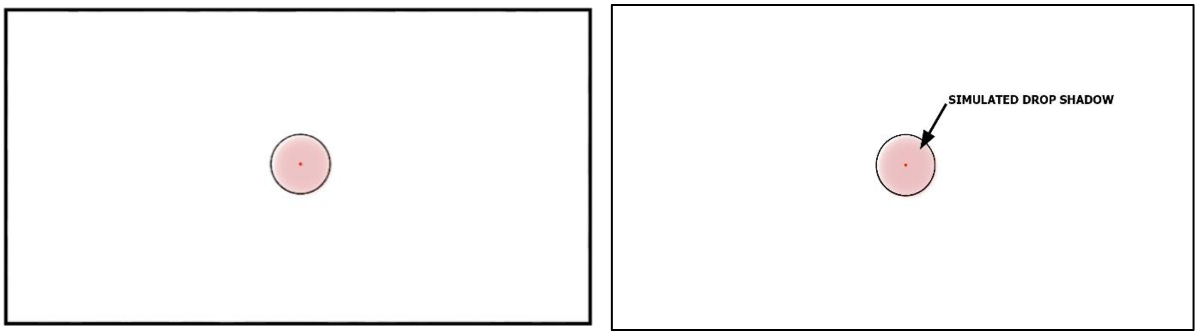
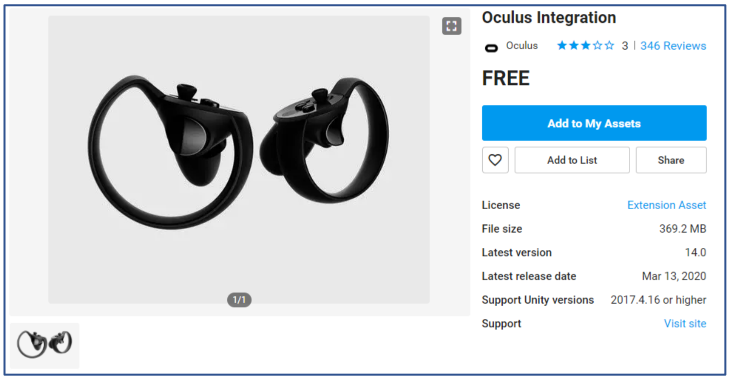


Figure 10. An image of what the researcher saw in the virtual HMD. A simulated afterimage has been applied in the form of a faded red circle for demonstration purposes. Note that the simulated afterimage failed to fill the calibration ring completely and would therefore require that the VR environment rendered by Unity would need to be scaled in order for the environment to become veridical.

3.2.5) Motion Parallax: In order to create motion parallax, it was necessary to track the Oculus Rift VR headset in all our Unity applications using the ‘Oculus Integration’ package provided by Oculus for Unity. While there is some basic head tracking already built into Unity for VR headsets, we required support for Oculus’s Touch Controllers in order to calibrate our scene. As such we required the ‘Oculus Integration’ for Touch Controller support (Figure 11).



Oculus Integration

Oculus ★★★☆☆ 3 | 346 Reviews

FREE

License	Extension Asset
File size	369.2 MB
Latest version	14.0
Latest release date	Mar 13, 2020
Support Unity versions	2017.4.16 or higher
Support	Visit site

Figure 11. An image of the Oculus Integration package provided by Oculus.

3.3) Headsets

For all Virtual conditions, stimuli were presented using an Oculus Rift CV1 virtual reality headset (Oculus VR, 2016). In order to equate all HMD-related influences from the Virtual conditions to that of the Real-World conditions, a mock HMD with dimensions and weight equated to the Oculus CV1 was used (Figure 12). The mock HMD contained two plastic lenses with no refractive power. The use of these “lenses served two purposes: 1) they introduced some of the peripheral chromatic aberration seen in the Oculus CV1 lenses, and 2) restricted the Field-of-View (FOV) to 100 degrees to match that of the CV1 headset. Lastly, an opaque, black, plastic “door” was attached to the front of the mock HMD, such that the “door” could be flipped down/up by the researcher to control target viewing time (Figure 13). The virtual scene also had a black “blind-closing” animation added in order to mimic the mock HMD “door-closing” movement induced by the researcher. Both of the Virtual HMD and the mock HMD were adapted from Lindsay’s experiment (Lockwood, 2017).

Figure 12. An image of Oculus Rift CV1 and the mock HMD developed for this experiment. Image adapted from Lockwood, 2017.

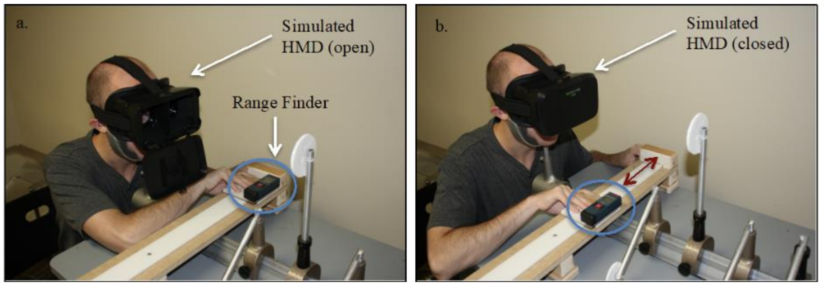


Figure 13. The image above demonstrates that the "door" on the mock HMD could flipped down/up by the researcher to control target viewing time. Image adapted from Lockwood, 2017.

3.3) Distance Measurements

Participant's distance estimations were recorded using an Optotrack motion tracking system (Optotrack Certus; sampling rate at 200 Hz; Figure 14). The system captured the three-dimensional, real-world positions of four Infrared-Emitting-Diodes (i.e. IREDs; See Figure 15) whose data were used to calculate three distance measurements in the current experiment. The three distance measurements calculated were: 1) distance between left and right index fingertips, 2) distance between the right-index fingertip to a point on the virtual reality HMD that was equivalent to eye distance from targets, and 3) distance between right-index fingertip to mock virtual reality HMD that was equivalent to eye distance from targets. Therefore, to calculate those three distance measurements, an IRED was attached to each of the participant's left and right index fingertips (Perceptual-Estimations), as well as an IRED placed on both the Oculus Rift CV1 and the mock VR headset (for Action-Estimations).

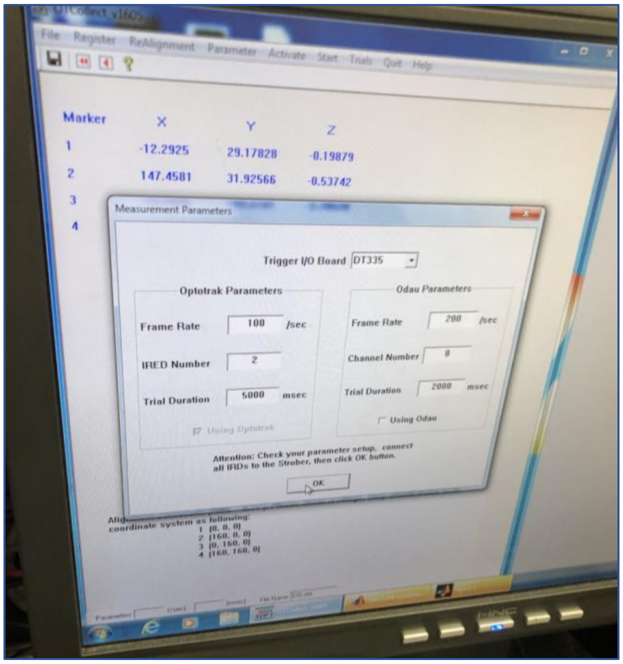
Figure 14. This is an image of the Optotrack motion tracking system.

Figure 15. An image of a pair of Infrared Emitting Diodes (IREDs).

4. Procedure

4.1) Preparation and Set-up

At the beginning of each testing day, an alignment file was collected for Optotrak Certus (frame rate 200/sec, IRED number 6; trail duration 7500 msec) to ensure that real-world coordinated of the tracking IREDs could be calculated for the physical space in which the experiment was being conducted (Figure 16). This process involves placing a flat calibration board, embedded with four IREDs at specific distances from one another, onto the experiment table at a specific position. The position of this calibration board and it's four IREDs is then recorded with the Optotrak system, the data of which is used to transform IRED position and distances into a real-world coordinate frame.



Marker	X	Y	Z
1	-12.2925	29.17828	-0.19879
2	147.4581	31.92566	-0.53742
3			
4			

Trigger I/O Board DT335

Optotrak Parameters
Frame Rate 100 /sec
IRED Number 2
Trial Duration 5000 msec
☑ Using Optotrak

Odau Parameters
Frame Rate 200 /sec
Channel Number 8
Trial Duration 2000 msec
☐ Using Odau

Attention: Check your parameter setup, connect all IRDs to the Strober, then click OK button.

Alig
coordinate system as following:
1 [0, 0, 0]
2 [160, 0, 0]
3 [0, 160, 0]
4 [160, 160, 0]

Figure 16. An image of Optotrack Alignment. Note: Parameters in the image are in default.

Due to the tracking system of the virtual HMD being setup each day (i.e. VR tracking cameras were not permanently mounted in the testing room), a calibration process was required prior to each testing day, so that the virtual room perfectly aligned to the real-world room. To use the calibration system, the researcher positioned themself at the real-world experiment table, then placed the Oculus Touch controllers against the two nearby corners of the table, thus landmarking the real-world coordinates of the experimental apparatus. Due to the fact that the VR controllers are accurately modeled and rendered in the VR scene, the researcher can see any misalignment between worlds as a space between the corners of the VR experimental table and the controllers. The researcher could then manipulate the VR controllers in such a way to allow the user to line up both front corners of the real-world table with that of the virtual table (See Appendix B for detailed calibration steps). This procedure, when coupled with the previously described After-image Paradigm, ensures that the virtual world is accurately scaled and positioned so as to be veridical.

4.2) Participant Screening

Upon arrival, participants were provided with a letter of information about the experimental procedure, including any associated benefits and risks, followed by an informed consent form. Following the signing of the informed consent, participants were asked to complete a brief questionnaire which included questions regarding demographic information and inclusion/exclusion criteria, as well as if there is any history of Epilepsy or seizure (See Appendix A). If participants did not have normal or corrected-to-normal acuity by way of contact lenses or corrective surgery, and/or had any history of seizures/epilepsy, they were excluded from testing. As was previously stated, flashing images from HMDs can cause seizures in persons suffering from photosensitive epilepsy (da Silva & Leal, 2017); therefore, poses a risk in some VR research. However, as the stimuli used in the current experiment did not contain any stroboscopic effects, the likelihood of inducing a seizure was remote. In order to eliminate all risk, any persons with a history of seizures were excluded from participating. In the event that a participant was excluded, they were provided with a debriefing form and were still rewarded with a participation credit.

After the completion of the questionnaire, a measure of participant’s stereoacuity was conducted using the RANDOT Stereotests (Stereo Optical Company Inc.). As discussed earlier, if stereoscopic acuity was found to be equal to or better than 40 seconds of arc, participants were allowed to proceed to the testing phase. Even if a participant did not pass the stereoacuity test, they were still allowed to participate in the experiment; however, their data were excluded from statistical analysis.

4.3) Testing

After the completion of all screening procedures, participants were randomly assigned to a pre-determined, randomized order of testing conditions. They were asked to sit in a chair with their chin either placed in the chin rest or slightly above the chin rest, depending on their first experimental condition. If their chin was placed in the chin rest, the height of the chin rest was adjusted until their eye height was equal to that of the center of the target stimuli. The chin rest was also moved either further or nearer to the target stimulus to ensure the distance was maintained as 60 cm in all testing sessions. Then, after receiving verbal consent from the participants, an IRED was tapped to each of the participants' left and right index fingertips and one IRED was tapped to each of the HMDs. Participants were then given the appropriate headset (i.e., VR HMD or Mock HMD) depending on their first condition. Once participants put on the headset and properly fit it to their head, the first block of testing began.

The following is a description for all testing conditions. The procedure for a single trail differed slightly based on the condition. Once participants were properly fitted with the condition's HMD, a series of tones were presented which served as indicators through each trial. At the onset of a trial, the participant heard three low tones and two high tones to indicate a target was about to appear. At the offset of a trail, the participant again heard three low tones and two high tones to indicate the target was about to go extinct and an estimation needed to be made upon hearing the last high tone (Figure 17).

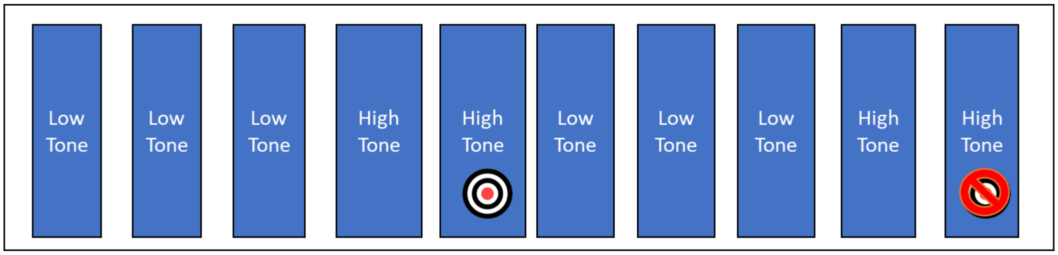


Figure 17. This image shows the onset/offset of tones, each tone was 500ms in length. At the second high-tone, target appears. At the last high-tone, scene turns dark and estimation made.

For all Real-World conditions, the front cover of the mock HMD was opened manually by the researcher to view the target. Then, after hearing the last high tone (Figure 17), the cover was closed by the researcher and participant made an estimation of egocentric distance. Similarly, for all VR conditions, the virtual stimuli appeared after hearing the second high tone, and then, the virtual display turned into dark after hearing the last high tone, at which point, participants made an estimation of egocentric distance.

For all Action-Estimation conditions, participants reached forward to the perceived location of the target (performed open-loop) with participant’s right index finger held at the perceived target distance. Similarly, for all Perception-Estimation conditions, participants used the distance between their index fingertips to indicate object distance.

For all the Head-Fixed conditions (i.e., zero Motion Parallax), participants were asked to place their chin in a chin rest, the placement of which physically restricted head movement and provided participants with head movement feedback, thus adding to the elimination of motion microparallax. To further ensure that there was no motion parallax information provided, the virtual reality system had head tracking turned off during these conditions, such that the view of the target remained centered and fixed in the HMD, regardless of head movement. For all Head

Non-Fixed conditions, the chin rest was lowered to allow the participant to freely move their head from side-to-side. In the VR condition, head tracking was turned on so that motion parallax was rendered in the VR HMD.

Once participants completed 15 tails in a given block (i.e., five repeats of each target distance), the headset was switched as needed, and the next block of trials would begin. After completion of eight randomized condition blocks, a final block consisting of nine calibration trails was conducted, allowing MATLAB to record the actual distances of targets within the space, the data of which would be used to ensure accuracy of distance measures.

Results

Due to the rising number of COVID-19 cases in 2020, it was recommended in early March, that "in-person" testing of participants be ceased at the Huron University and University of Western Ontario. Unfortunately, this timing coincided with the onset of data collection for this project. As a result, only one participant was tested with the full set of conditions in the final experimental configuration. In an effort to explore the data, this full dataset was combined with the partial dataset (Head Non-Fixed conditions only) of a final pilot subject. Therefore, data for only half of the experimental conditions was collected (Head Non-Fixed conditions only) and used in data analysis (N = 2). These data were analyzed using a three factor, within-subjects ANOVA with two levels of Environment (Real-World versus Virtual Reality), two levels of Estimation Type (Perception-Estimation versus Action-Estimation), and three levels of Target Distance (30 cm, 45 cm, 60 cm).

Figure 18 represents participant's averaged performance across conditions, which clearly illustrates that there are likely no significant main effects of Target Distance, Environment, nor

Estimation Type. Likewise, no significant two-way interactions between Target Distance and Environment, Target Distance and Estimation type, and Environment and Estimation Type appears likely, nor does a three-way interaction between Target Distance, Environment and Estimation type. These observations are clearly supported by the statistical analysis results provided in Table 1.

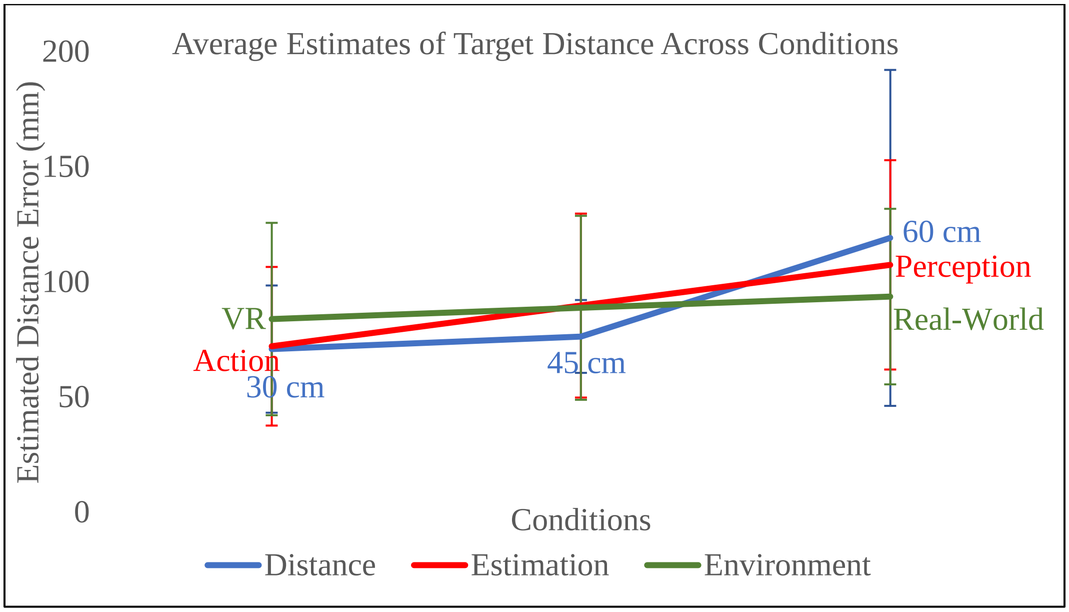
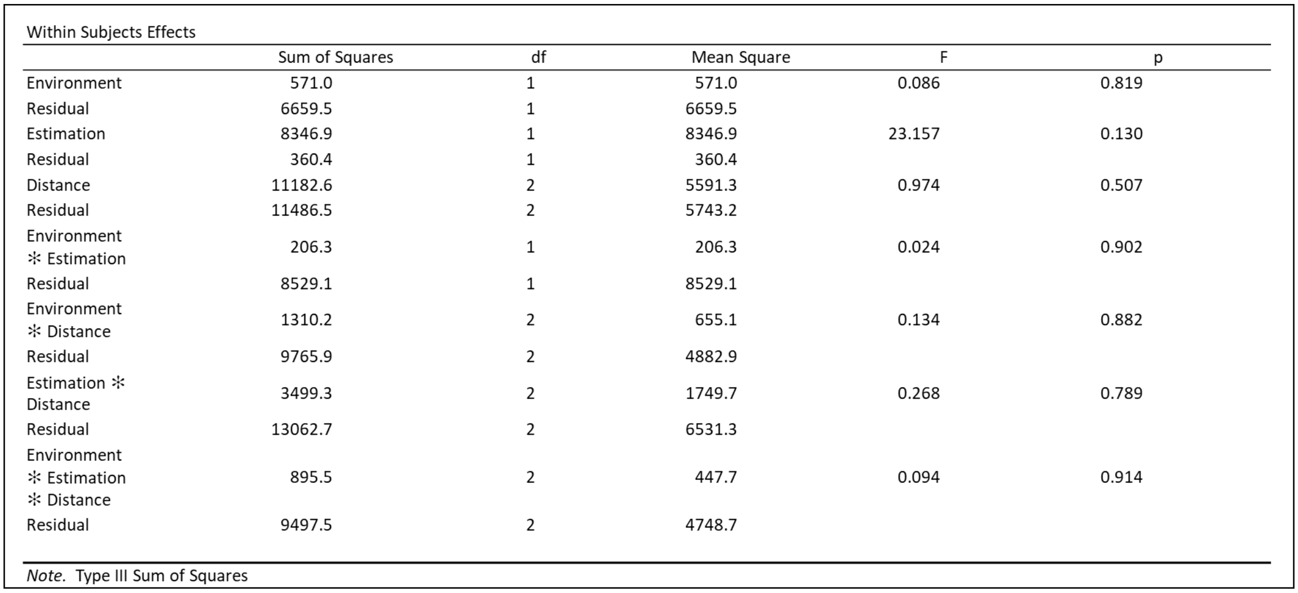


Figure 18. The figure above illustrates average estimates of target distances across the three condition factors: Target Distance, Environment and Estimation Type. NOTE: Error bars presented above are Standard Error of the Mean.



Within Subjects Effects

	Sum of Squares	df	Mean Square	F	p
Environment	571.0	1	571.0	0.086	0.819
Residual	6659.5	1	6659.5		
Estimation	8346.9	1	8346.9	23.157	0.130
Residual	360.4	1	360.4		
Distance	11182.6	2	5591.3	0.974	0.507
Residual	11486.5	2	5743.2		
Environment ✻ Estimation	206.3	1	206.3	0.024	0.902
Residual	8529.1	1	8529.1		
Environment ✻ Distance	1310.2	2	655.1	0.134	0.882
Residual	9765.9	2	4882.9		
Estimation ✻ Distance	3499.3	2	1749.7	0.268	0.789
Residual	13062.7	2	6531.3		
Environment ✻ Estimation ✻ Distance	895.5	2	447.7	0.094	0.914
Residual	9497.5	2	4748.7		

Note. Type III Sum of Squares

Table 1. This table illustrates the statistical results from the 2 x 2 x 3 within-subjects ANOVA.

Discussion

From gaming industries to training health-care professionals, the application of Virtual Reality (VR) has become more and more practical, affordable and convenient in recent years. With this raising interest in the investigation of using VR technology as a training tool, recent research has suggested that current VR hardware is still plagued by distance misestimation errors (Armbürster et al., 2008; Narceri et al., 2011).

The goal of the current experiment was to investigate whether the addition of Motion Parallax would allow participants to make more accurate distance estimations, in both the real and virtual worlds, as well as to determine whether Perception- and Action-Estimations were affected similarly. It was hypothesized that the addition of Motion Parallax would make participants' distance estimations more accurate in VR as compared to when Motion Parallax was not present. Further, based on past findings (Lockwood, 2007), it was hypothesized that participant would be more likely to underestimate distance when using Perception-Estimation than using Action-

Estimation that distance-estimation errors would increase as the target distance increased, and that Real-world estimates would be more accurate than those if the virtual-world.

Expected Results

Due to rising number of COVID-19 cases in 2020, all in-person testing needed to cease and a full array of testing conditions on twenty-four participants could not be completed. As a result, only one participant was tested with the full set of conditions in the final experimental configuration. In an effort to explore the data, this full dataset was combined with the partial dataset (Head Non-Fixed conditions only) of a final pilot subject. Therefore, data for only half of the experimental conditions was (Head Non-Fixed conditions) used in data analysis (N = 2), which resulted in low statistical power. Likely leading to why no significant main effects, nor interactions were found.

Figure 19 illustrates the expected results if the full array of testing conditions on twenty-four participants were able to be collected. Clearly, the results indicate a main effect of distance, which suggests that estimation error increases as the target distance increases. A main effect of environment type, which demonstrates that participants perform better in Real-world than VR. A main effect of estimation type, which shows participants are ore likely to underestimate when using Perception-Estimation than Action-Estimation. Also, with the addition of Motion Parallax, less distance-estimation errors should be observed across all conditions.

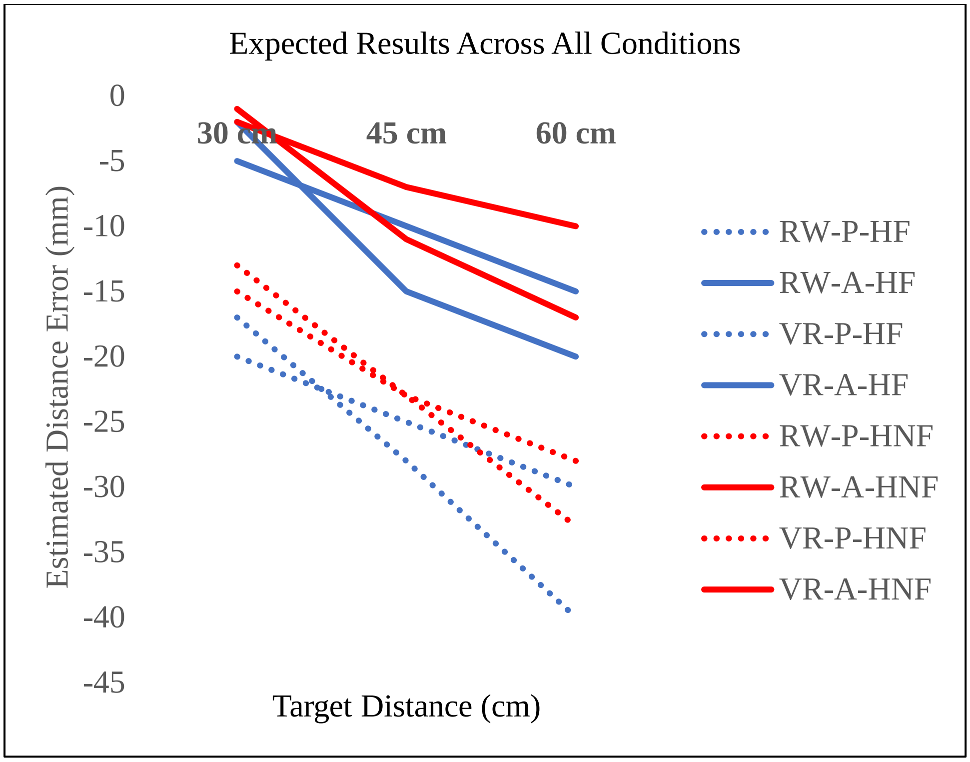


Figure 19. Expected results across all conditions with red lines indicating Head Non-Fixed conditions; whereas, blue lines indicating Head-Fixed conditions. Solid lines represent Action-Estimation conditions and round-dot lines represent Perception-Estimation conditions. Error bars represent the standard error.

Limitations

Typically, there is a section discussing limitations of the project within a thesis. However, for the current experiment, we invested a great deal of time and effort into building testing stimuli and to developing a method for spatially matching the environmental dimensions between the Real-World and VR. Specifically, as discussed in the Method section: 1) all of the retinal images of the “bull’s-eye” targets subtended the same visual degrees so that object size could not be used as a cue to distance, 2) ensured that the scale of the virtual scene was accurately recreated within Unity, 3) ensured that the perceived size of objects was accurate when viewed through the

HMD by creating an After-Image Paradigm, which compares and contrasts the afterimages of real targets to ensure size accuracy in VR, and 4) all of our distance estimations were recorded using an Optotrack motion tracking system, which has a positional accuracy of up to 0.1 mm and resolution of 0.01 mm, the use of which would reduce any measurement errors that may be present when manual measurement methods are employed.

Future Direction

Once the COVID-19 pandemic has subsidised, I intend on returning to the lab in order to fully collect data from all twenty-four participants and complete the current project. If I discover that with the inclusion of Motion Parallax does not eliminate the spatial representation errors, as we suspect, the lab intends to investigate the role of the conflict within the Accommodation Reflex that exists when using a VR HMD. This reflex links vergence eye position and accommodative state in such a way that they function in an orchestrated fashion and normally work together to calculate accurate object distance estimates (Emslie, Sachs, Claassens, & Walters, 2007). However, due to the fact that all previously and currently manufactured VR HMDs only utilize vergence eye position changes to drive a perception of depth, while maintaining a constant accommodative state, this pits the two depth cues of the Accommodative Reflex against one another. As vergence eye position and accommodative state are both used in the calculation of foveated object distance, forcing the individual to accommodate to a fixed distance should introduce error into distance estimates. This conflict is easily seen when viewing a 3D movie and you notice that you are having some difficulty focusing on the images presented, as your eyes verge to one distance, but are accommodating to another (Keller & Colucci, 1998). This is especially noticeable as you switch from verging to one virtual distance to another and you have some difficulty maintaining focus. The ideal VR HMD would not only use vergence eye position cues, but also alter the HMD optics such that the lenses of viewer's eyes would need to accommodate appropriately to the verged distance. This would eliminate the conflict between these two depth cues of the Accommodative Reflex and should result in more accurate perceptions of depth. In fact, my supervisor (Dr. Derek Quinlan) is currently working on an HMD prototype that would address this very issue and plans to conduct a study to determine whether distance estimates become more accurate once these two linked depth cues are brought into agreement.

References

Aggarwal, R., Black, S. A., Hance, J. R., Darzi, A., & Cheshire, N. J. W. (2006). Virtual reality simulation training can improve inexperienced surgeons' endovascular skills. *European Journal of Vascular & Endovascular Surgery, 31*(6), 588-593. doi:10.1016/j.ejvs.2005.11.009

Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B. (2008). Depth perception in virtual reality: Distance estimations in peri- and extrapersonal space. *CyberPsychology & Behavior, 11*(1), 9-15. doi:10.1089/cpb.2007.9935

Dayal, R., Faries, P. L., Lin, S. C., Bernheim, J., Hollenbeck, S., DeRubertis, B., . . . Kent, K. C. (2004). Computer simulation as a component of catheter-based training. *Journal of Vascular Surgery, 40*(6), 1112-1117. doi:10.1016/j.jvs.2004.09.028

Dmitrii Tiron, Michael Langer; Microparallax is preferred over blur as a cue to depth order at occlusion boundaries. *Journal of Vision 2018;18*(10):491. doi: https://doi.org/10.1167/18.10.491.

Emslie, R., Sachs, N., Claassens, A., & Walters, I. (2007). The near triad and associated visual problems. *African Vision and Eye Health, 66*(4), 184-191. doi:10.4102/aveh.v66i4.256

Goldstein, E. B. (2013). *Sensation and perception* (tenth ed.). Belmont, California: Wadsworth.

Goodale, M. A., & Milner, A. D. (1992). *Separate visual pathways for perception and action.* England: Elsevier Ltd. doi:10.1016/0166-2236(92)90344-8F

Hoffman, D. M., Girshick, A. R., Akeley, K., & Banks, M. S. (2008). Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of Vision, 8*(3), 33.1. doi:10.1167/8.3.33

Keller, K., & Colucci, D. (1998). Perception in HMDs: What is it in head-mounted displays (HMDs) that really make them all so terrible? Paper presented at the, *3362*(1) 46-53. doi:10.1117/12.317454

Kongsilp, S., & Dailey, M. (2017). Motion parallax from head movement enhances stereoscopic displays by improving presence and decreasing visual fatigue. *Displays, 49*(C), 72–79. https://doi.org/10.1016/j.displa.2017.07.001

Lockwood, L. (2017) Depth Perception in Virtual Peripersonal Space: An Investigatioon of Perceptual- VS Action-Based Estimates. London, Ontario: Huron University College.

Murgia, A., & Sharkey, P. M. (2009). Estimation of Distances in Virtual Environments Using Size Constancy. *International Journal of Virtual Reality, 8*(1), 67-74. https://doi.org/10.20870/IJVR.2009.8.1.2714

Naceri, A., & Chellali, R. (2011). Depth perception within peripersonal space using head-mounted display. *Presence: Teleoperators and Virtual Environments, 20*(3), 254-272. doi:10.1162/PRES_a_00048

Napieralski, P., Altenhoff, B., Bertrand, J., Long, L., Babu, S., Pagano, C., . . . Davis, T. (2011). Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception (TAP), 8*(3), 1-19. doi:10.1145/2010325.2010328

Oculus Rift (2019, December). *Initiatives.* Retrieved from https://www.oculus.com/rift/

Parmeggiani, P., & Parmeggiani, P. (2016). From grand tour to virtual tour: Italy through the stereoscope in 1900. *Visual Studies, 31*(3), 231-247. doi:10.1080/1472586X.2016.1209985

Peters T.M., Linte C.A., Moore J., Bainbridge D., Jones D.L., Guiraudon G.M. (2008) Towards a Medical Virtual Reality Environment for Minimally Invasive Cardiac Surgery. *Medical Imag. Augment. Reality 5128*, 1–11.

Seymour, N. 2008. VR to OR: A review of the evidence that virtual reality simulation improves operating room performance. *World J. Surgery 32*, 2, 182–188.

Sutherland, I. 1965. The ultimate display. *In Proceedings of the IFIP Congress*. Vol. 2, 506–508.

Appendix A

ELIGIBILITY QUESTIONAIRE

Basic information

Age: _____

Gender (Please Circle): M F Identified as______

Eligibility information (please circle)

1. Do you have normal, or corrected to normal vision? (20/20)

   Yes No

2. If you have corrected vision, please indicated which of the following applies to you at this moment.

   I am wearing contact lenses I had corrective surgery I am wearing glasses

3. Are you right-handed?

   Yes No Ambidextrous

4. Do you have epilepsy? Or, have you had a seizure?

   Yes No

5. Have you had severe side effects, such as nausea, vomiting from exposure to virtual reality in the past?

   Yes No

Appendix B

CALIBRATION STEPS

1) When the application loads, usually the view is way above the floor in the room.

2) The researcher then positioned himself in the chair in front of the table being used for the study. The researcher placed the right-hand Oculus Touch Controller against the nearest right table corner. He pressed the inner trigger of the right-hand controller to set the height and position of the room.

3) Next, the researcher rotated his right wrist on the yaw axis to change the orientation of the room until he has aligned the left-hand Touch Controller with the nearest left table corner.

Result after fixing the orientation:

4) After the room has been calibrated, the settings can be locked by pressing one of the right-hand controller's face buttons.

(The end of calibration)

Curriculum Vitae

Name: Hongyao Zhu

Place and Year of Birth: Xi'an, China, 1998

Secondary School Diploma: Sunshine Middle School, Xi'an, China

Awards: Huron's International Scholarship Award $ 3,000.00 (2016)