Spring 5-7-2022

Increasing Perceived Realism of Objects in a Mixed Reality Environment Using 'Diminished Virtual Reality'

Logan Scott Parker

University of Mississippi

INCREASING PERCEIVED REALISM OF OBJECTS IN A MIXED REALITY ENVIRONMENT USING “DIMINISHED VIRTUAL REALITY”

by
Logan Scott Parker

A thesis submitted to the faculty of The University of Mississippi in partial fulfillment of the requirements of the Sally McDonnell Barksdale Honors College.

Oxford
May 2022

Approved by

____
Advisor: Professor J. Adam Jones PhD

____
Reader: Professor Charles Walter PhD

____
Reader: Professor Yixin Chen PhD

ACKNOWLEDGEMENTS

I would like to thank my thesis advisor, Dr. J. Adam Jones, who has a passion for education and knowledge that is unparalleled. He is a true scientist through and through. His dedication and true joy for data collection, manipulation, and visualization has inspired countless up and coming researchers and professionals. I would also like to thank Dr. Charlie Walter for his support and guidance throughout the entire process of completing this research. He was always there to critique and applaud the work I was doing along the way. I would also like to acknowledge the University of Mississippi and specifically the Department of Computer Science for continually pushing me to become a better student and citizen scholar. Finally, I would like to thank the Sally McDonnell Barksdale Honors College for everything they do to foster growth academically and personally throughout every member's time at the University and beyond.

ABSTRACT

With the recent explosion of popularity of virtual and mixed reality, an important question has arisen: "Is there a way to create a better blend of real and virtual worlds in a mixed reality experience?" This research attempts to determine whether a visual filter can be created and applied to virtual objects to better convince the brain into interpreting a composite of virtual and real views as one seamless view. The results found in this study show that when presented with a scene composed of a combination of visually similar stimuli where one is virtual and the rest are not, participants of the study could not reliably identify whether all objects were real or if some of the objects were virtual after being exposed to the stimulus for 300ms. The "eccentricity effect" is an effect observed by Carrasco et al. [1] which describes the phenomenon where visual search tends to be better (faster and more accurate) when the target object is more central to the fovea and worsens as optical eccentricity increases. This effect was not present in the data from this study as a pronounced effect. The proposed factors that contributed to these results are further outlined and discussed in this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The Oculus Quest 2, the most ubiquitous VR headset in the world [2], has recently been updated with the ability for developers to build mixed reality experiences using this relatively low-cost head mounted display (HMD). The Oculus Quest 2 is a virtual reality headset first and foremost, but it utilizes four cameras embedded into the device to achieve inside-out tracking. This means that no external tracking is required for the device to function. These cameras capture the physical room surrounding the HMD in grayscale with a high amount of visual noise.

The Oxford Languages defines mixed reality as, "A medium consisting of immersive computer-generated environments in which elements of a physical and virtual environment are combined." Because the cameras on the Quest 2 are low resolution, grayscale, and grainy, they were most likely never intended to be used for mixed reality experiences; however, they have recently been updated with the software support to enable mixed reality development. Many developers have already begun creating exciting mixed reality experiences to take advantage of this new feature, but a problem exists with the way in which mixed reality experiences are implemented in the Oculus Quest 2. The visual fidelity of virtual objects rendered by the HMD and superimposed onto the camera feed far exceeds the visual fidelity of the grayscale, noisy camera feed. This creates a break in immersion and a lapse in the visual cohesiveness of the experience. This thesis aims to close that gap in immersion and create a more consistent and believable mixed reality experience for users.

This experiment is effectively a combination of previous research done in an area called diminished reality. Diminished reality is a term for the method of blending the real and virtual (computer drawn) world in which the real-world view is diminished to a lower fidelity sample in order to hide unwanted portions of the real world or match the appearance of the virtual world [3]. One method commonly used to evaluate diminished reality is the utilization of visual-search tasks. Visual-search tasks are implemented in VR by flashing an image for a few hundred milliseconds and evaluating a participant's ability to visually search the image and find a given stimuli among distractors or visually similar elements.

This research implemented the use of an Oculus Quest 2 (now renamed the Meta Quest 2). The Quest 2 was chosen as the primary platform because of its low cost and ubiquity. When considering the userbase of virtual reality headsets, as of January 2022, the Quest 2 had a 46% market share of all VR headsets on Steam, the most widely used game distribution service on PC's [2]. This measure of market share does not even consider the users of the Quest 2 who have not interacted with steam. This means that the research conducted on this headset will be quite applicable to the modern VR market.

The Quest 2 has four integrated cameras on the exterior of the device. These cameras can be used to inject a view of the physical room that the headset is in into the view of the wearer of the headset enabling video see-through augmented reality. Oculus calls this feature "Passthrough" [4] because the real world is being passed through the headset from the cameras on the exterior to the user of the headset's eyes. On the product page for the headset, Oculus states, “Passthrough uses the sensors on your headset to approximate what you would see if you were able to look directly through the front of your headset and into the real world around you.” [4] The passthrough cameras, while full of potential for creating mixed reality experiences in a much cheaper package than alternative headsets, are limited to a grainy, grayscale view of the real world. It was this limitation, however, as well as the recent explosion in popularity of the Oculus Quest 2 which inspired this research.

Previous experiments in diminished reality have always diminished the real-world camera view to match the lower visual fidelity produced by virtual rendering [3, 5]. This work, however, flips that idea and attempts to diminish the quality of the virtual world to match the low fidelity, grayscale cameras incorporated into the Oculus Quest 2. The name we have decided to call this method by is *diminished virtual reality*. There is extensive research into how to bring the virtual world into a higher fidelity, but in this case, it is not necessary to create a wholly convincing mixed reality experience. This paper hypothesizes that when the qualities of the real camera view and the virtual objects match closely enough, the user perceives the combination of these two layers as a cohesive environment.

In my previous development of mixed reality experiences for the Quest 2, it has been quite detracting from the overall experience and believability when the virtual objects being rendered and superimposed into the real-world view are of a much higher fidelity and are in full-color while the passthrough cameras are capturing low-fidelity, grayscale view. This quirk that comes from developing mixed reality experiences on the most widely used MR capable headset [2] led to the development of a shader which can be applied to virtual objects in an attempt to mimic the color-grading, textural noise, and blurriness created by the passthrough cameras.

CHAPTER 2: RELATED WORK

In 1994, Steve Mann outlined the first example of diminished reality as a use for ‘reality mediators,’ a term Mann coined to encapsulate the entire genre of devices which altered the real view of a user [5]. He describes reality mediators as devices which use video cameras affixed to an HMD with displays. The cameras pass through the video feed to the displays after applying some computation to it. This describes the implementation of Oculus “Passthrough” very well. Mann proposed that diminished reality could take the form of diminishing the quality of the real world by reducing the color space to let users experience color blindness or reducing the resolution of the real world by applying the appropriate “visual filter” to create the desired degree of degradation. Mann’s proposed form of diminished reality and all other previous work in the field of diminished reality has focused on removing, masking, or diminishing real objects from the synthesized mixed reality view [5, 3]. In our *diminished virtual reality* approach, the reality mediator applies appropriate ‘visual filters’, in this case shaders, to existing 3D models to diminish both the color space and the resolution of the virtual objects superimposed into the scene. Shaders are used in the rendering of computer graphics to produce desired levels of lighting, color grading, and overall visual appearance. This is important because it has become commonplace to find hyper realistic 3D models for use in VR applications, but these models do not match the view of the real world produced by the cameras of the Quest 2. Mann created an example of one of these reality mediators which he used to achieve diminished reality. In his paper *Mediated Reality*, Mann said, “Unlike typical beam-splitter implementations of augmented reality, transparency, if desired, is synthesized, and therefore is only as good as the components used to make the RM.” [5] It is incredible that the problems Mann was outlining nearly 30 years ago are still relevant today. The Quest 2 captures the real world in grayscale, so the transparency of the headset is limited to grayscale.

This thesis also aimed to discover if there is a relationship between virtual object ocular eccentricity and a user’s ability to detect a virtual object that had been diminished to resemble the grayscale, low-fidelity camera feed produced by the Quest 2 cameras. Previous research by Carrasco et al. [1] has attempted to quantify the effects of target eccentricity on the efficiency and accuracy of search performance in a visual search task. In a visual search task, a participant is to exposed to an array of stimuli which the participant must search for a specific object or target. Carrasco et al. has shown, through visual search tasks, that targets presented near fixation are, in fact, found more efficiently than are targets presented at more peripheral locations [1]. Carrasco et al. called this effect the “eccentricity effect”. The research conducted by Carrasco et al, however, utilized target stimuli which were salient from their background and non-target stimuli. This thesis aimed to implement the methods for evaluating the effects of eccentricity on visual search for target stimuli which were non-distinct from both their background and non-target stimuli.

Previous work from Wolfe et al. [6] attempted to determine whether the primary cause of the "eccentricity effect" defined by Carrasco et al. is an attentional bias that allocates attention preferentially to central items. They found that peripheral reduction in visual sensitivity, peripheral crowding, nor cortical magnification was responsible for the eccentricity effect. They conclude that an attentional bias causes participants of visual search tasks to search from fixation outward.

Aside from target eccentricity, the effects on response time and accuracy in visual search tasks caused by non-target and target similarity has been examined in research by Duncan and Humphreys [7]. Their research found that error rates and reaction times increased with increasing target-non-target similarity. One significant trend found by Duncan and Humphreys was that as the number of nontargets increased, so did error rate and response times. Their research, however, did not consider similarity between background and targets/non-targets.

A. S. Hess et al. extended on Duncan and Humphreys research in the paper *On the Hunt: Searching for Poorly Defined Camouflaged Targets* [8]. Hess et al. employed an image of nature with the target being a section of the image which had been blurred. This created a target which was highly similar to non-targets. Hess et al. found that the larger the target, the higher the accuracy and the lower the reaction times. This thesis is similar to Hess's work because of the similarity between non-target and target stimuli, but Hess et al. did not evaluate the effects of the degree of similarity between targets and non-targets, only the size of the target.

M. Neider and G. Zelinsky evaluated the effects of target-background similarity in their paper *Searching for Camouflaged Targets: Effects of Target-Background Similarity on Visual Search [9]*. They found that as target-background similarity increased, so did error rate and response times. They noted, however, that eye movement analyses revealed a biased search process which resulted in participants preferentially fixating on discrete distractors rather than the target-similar background. Their research shows that even in scenes with high target-background similarity, “salient patterns segmented from a background are preferred while more target-similar unsegmented regions of the background are relatively neglected” [9].

The typical exposure time for these visual search tasks range from 62ms to until the participant responded [1, 10]. In *Speed of Processing in the Human Visual System* by Thorpe et al. [11], previously unseen images were flashed on for 20ms and participants were able to recognize whether the image contained an animal. According to Thorpe et al., the visual processing time needed to perform the highly demanding task of processing a complex image is under 150ms. The results from Thorpe et al. were used as a basis for the setting the stimulus exposure time in the experiment conducted for this paper.

Target-background similarity [9, 8] as well as target-non-target similarity [7, 10] has been studied previously, but there has not been any research, to our knowledge, on visual search tasks in which targets are intentionally camouflaged by being similar to both background and non-targets. As such, the role that target eccentricity plays in this unique setting is also unstudied.

CHAPTER 3: APPLICATION

In order to conduct the trials for the experiment, an application first needed to be developed which would subject the participants to several trials with changing stimuli and record data about each trial. This application was developed using Unity 3D and utilized the latest version (37.0) of the Oculus Integration Package. On deployment, the application was packaged as an APK file to run on the modified version of the Android operating system that the Oculus Quest 2 uses. The Oculus Integration Package aids in the development of software which interacts with the proprietary Oculus software and hardware. The application which was developed can be broken apart into three distinct parts: a data manager script to parse and save incoming data, a scene manager script to manage when to present distinct configurations of the stimuli, and the Unity scriptable render pipeline to handle to rendering of the 3D scene.

The data manager captures the data corresponding to each trial and associates it to a participant using a participant id. As shown in figure 1, this data includes participant response for a specific trial, trial id, stimulus position relative to other distractors, stimulus coordinates, head coordinates, head distance from the stimulus, target eccentricity, response time, and whether the participant gave the correct answer. The data manager script also handled saving the data file to the internal storage of the headset in a CSV format. Using a CSV format allows for easy analysis in Excel or other statistical analytics software such as Python or R.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	Scene	SubjectId	Ball Postion	Ball X	Ball Y	Ball Z	Head X	Head Y	Head Z	SeatDistance	Eccentricity	Response	ResponseTime	Correct
2	1	13	3	-0.3428979	0.7588845	1.58698	-0.5610911	1.513205	0.4622548	1m	22.5	1	3.661949	0
3	2	13	0	-0.9418087	0.7588845	1.623117	-0.5698751	1.506809	0.4958292	1m	-45	0	1.268349	1
4	3	13	-1	-1.530738	0.7588845	1.658653	-0.568702	1.505627	0.5052638	1m	-1	1	1.184601	1
5	4	13	1	-0.7421718	0.7588845	1.611072	-0.5759842	1.50266	0.5141084	1m	-22.5	0	1.058723	1
6	5	13	4	-0.143261	0.7588845	1.574934	-0.5586709	1.504073	0.4789627	1m	45	0	1.295395	1
7	6	13	2	-0.5425348	0.7588845	1.599026	-0.5749393	1.499358	0.4887014	1m	0	1	1.587669	0
8	7	13	4	-0.143261	0.7588845	1.574934	-0.5664928	1.504193	0.4805199	1m	45	1	0.8915558	0
9	8	13	3	-0.3428979	0.7588845	1.58698	-0.5671724	1.501614	0.50766	1m	22.5	1	1.016815	0
10	9	13	0	-0.9418087	0.7588845	1.623117	-0.5644543	1.495601	0.5161968	1m	-45	0	0.91716	1
11	10	13	-1	-1.530738	0.7588845	1.658653	-0.5596647	1.496242	0.4806041	1m	-1	1	0.7808228	1
12	11	13	1	-0.7421718	0.7588845	1.611072	-0.562363	1.49747	0.5112377	1m	-22.5	0	0.8636322	1
13	12	13	2	-0.5425348	0.7588845	1.599026	-0.5570591	1.494394	0.5318416	1m	0	0	0.7527466	1
14	13	13	1	-0.7421718	0.7588845	1.611072	-0.5755742	1.497784	0.5321844	1m	-22.5	1	2.269302	0
15	14	13	0	-0.9418087	0.7588845	1.623117	-0.5695738	1.498603	0.5597427	1m	-45	0	0.782547	1
16	15	13	4	-0.143261	0.7588845	1.574934	-0.563781	1.501807	0.5087055	1m	45	0	1.490143	1
17	16	13	2	-0.5425348	0.7588845	1.599026	-0.5949417	1.497792	0.4634808	1m	0	1	1.40683	0
18	17	13	-1	-1.530738	0.7588845	1.658653	-0.5731005	1.501109	0.5127968	1m	-1	1	0.8208008	1
19	18	13	3	-0.3428979	0.7588845	1.58698	-0.568463	1.499852	0.5032183	1m	22.5	0	0.8215942	1

Figure 1

This figure shows some sample data collected by the data manager script.

The scene manager portion of the application is responsible for allowing the primary investigator (PI) to co-locate the real world with the virtual one. It also manages when to present a blank screen and when to expose the stimulus. This is all done using coroutines in Unity which allow for tasks to be asynchronously distributed among several frames rather than pausing execution of scene rendering. The most important task of the scene manager portion of the application is to choose a scene from a precomputed random shuffle and present them to the user. This order is available for the primary investigator to see so that the PI can change the real-world view to match what the software is expecting. This is a liability, however, because if the order of scenes the application is presenting is different from what the PI is arranging in the physical world, it could corrupt the data. To combat this, error checks have been implemented into this portion of the application which prevented the participant from continuing to the next scene if they have not yet provided a response. The headset plays an error sound audible to both the PI and the participant if the participant attempts to continue before responding. This allows both the PI and the participant to realize that a mistake was attempted. A second error case is also handled if the participant attempts to answer the question twice. In this case, the headset also plays an error sound and only records the first response. Both of these error checks ensure that the PI and the application stay synchronized.

The final portion of the application is the scene rendering. This portion includes the generation of the *diminished virtual reality* shader, outlined in section 4.2, that is responsible for adding color grading and noise to the rendered objects. The rendering portion is able to manipulate the size, location, rotation, lighting and other visual characteristics of the stimulus. The renderer also handles post-processing effects which enhances the overall blending of the scene by applying a general color-grading effect to the entire rendered scene.

CHAPTER 4: EXPERIMENT

4.1 STUDY DESIGN

The experiment was designed with 6 trials repeated 3 times in a predefined random order for a total of 18 trials per participant. The independent variable throughout the six trials was horizontal optical eccentricity. The horizontal visual eccentricities tested spanned a 45-degree section of the participants' field of view (FOV). This represents a span from -22.5 degrees from the participant's focal point to +22.5 degrees. This 45-degree field was divided into 5 distinct locations -22.5, -11.25, 0, 11.25, and 22.5 degrees. A single trial of the experiment consisted of 5 black balls being placed in a line on a table. Each ball position was a distinct location in the 45-degree span of the participant's vision. In each trial, one of the physical balls was either removed from the table in one of 5 possible eccentricity positions or none of the physical objects were removed. The physical ball that was removed was replaced by a co-located virtual version of the ball with the *diminished virtual reality* shader applied.

Participants of the study were asked to stand on a pre-marked location on the floor in front of the table with the stimuli; this can be seen in figure 2. To mark this standing location, a secondary mark was made on the floor directly below the table where the 0-degree ball would sit, then a distance of 1 meter was measured from this secondary location. This made the final standing position 1 meter from the location of the floor directly below the center stimulus. This ensured that the balls were at the correct eccentricities in the participant's FOV. Participants, wearing the Oculus Quest 2 head mounted display (HMD), were then presented with a black screen except for a fixation cross at the proposed 0-degree position. The participants were asked to stare at the fixation cross and when ready click a button they had been shown on the right Oculus Touch Controller. After the participant clicked this button, there was a 3 second delay before the participant was exposed to the passthrough view of the real-world table for 300ms.

The experiment was conducted using the Two-Alternative Forced-Choice paradigm. This paradigm describes a task in which participants are exposed to some stimuli and asked to answer some question with only two possible responses. The pattern of the participant's choices and the response times create a model of sensitivity to those stimuli. After being exposed to the passthrough view combined with the virtual objects using the *diminished virtual reality* shader, each participant was asked to answer the question "Were all of the objects on the table real?" The only responses that could be given were "all objects were real" or "at least one of the objects was virtual." The responses were recorded using the left and right Oculus Touch Controllers. The right-hand trigger indicated real; participants were told to remember "right all real." The left-hand trigger was used to indicate the alternative choice.

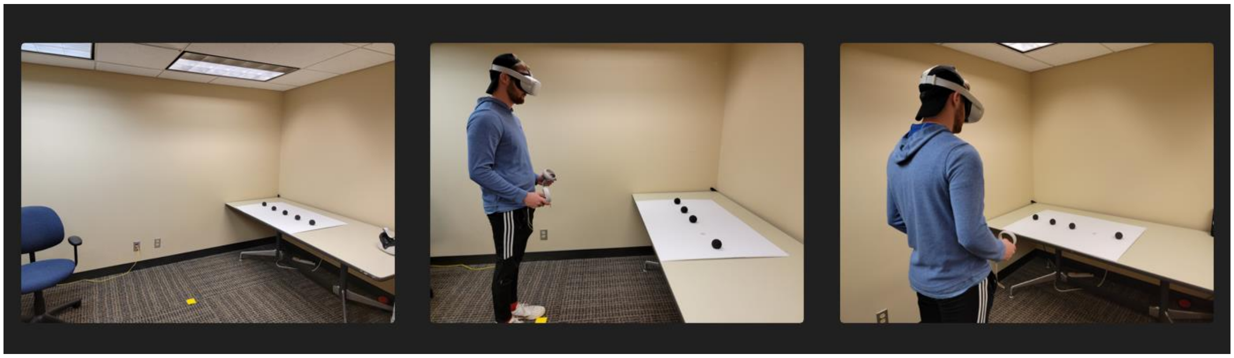

Figure 2

The experiment room and two images of how the experiment setup looked with a participant.

4.2 STIMULUS

The stimulus environment, shown in figure 3 (right), consisted of a passthrough view of the test environment, shown in figure 3 (left), and a virtual version of the missing ball overlaying the empty space where the real ball should have been. The virtual stimulus utilized the *diminished virtual reality* shader that was created using Unity Shader Graph shown in Figure 4. Unity Shader Graph is a tool that simplifies the creation of HLSL shaders by utilizing visual components and relationships rather than pure shader code. The shader created for this research subsampled the texture of the object into a grayscale color space and added in noise to mimic the noise generated by the passthrough cameras of the Oculus Quest 2. This was accomplished by using an oscillating sine wave with a remapped output from -1 through 1 to a larger range of 6 through 946. This larger range of sweeping numbers was used to control the offset for a Voronoi noise function. This noise function was then applied to the grayscale subsample of the texture of the virtual object using an overlay blend. The parameters on each of these functions, while generally arbitrary, were tuned through an iterative process to create the most convincing rendition of the effect created by the passthrough cameras.

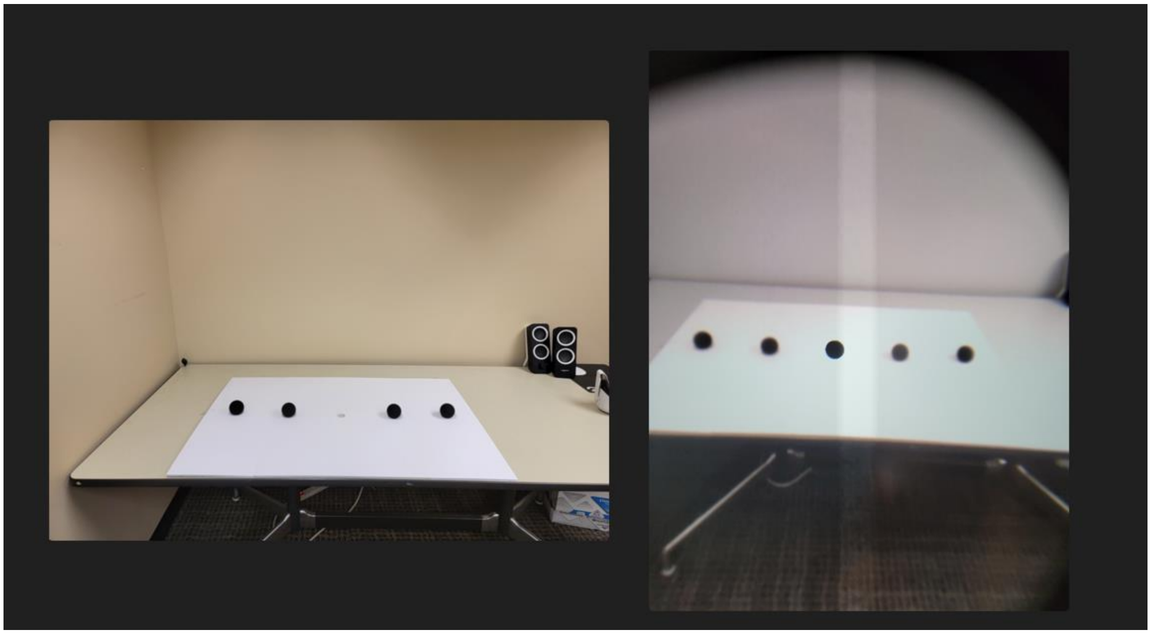

Figure 3

An image of the physical room (left) and an image of how the mixed reality scene appeared through the headset (right). The vertical line is an artifact of taking a photo through the lens of the HMD and was not visible to participants.

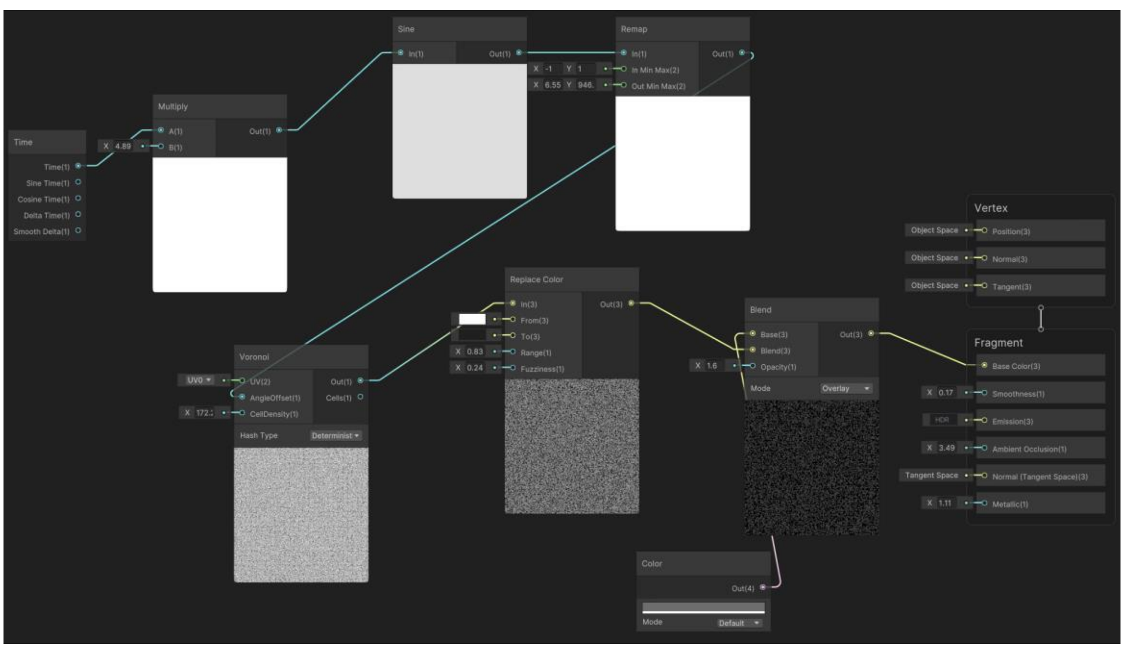

Figure 4

The Unity Shader Graph Structure which was iteratively developed to generate the color sampling and noise effects created by the passthrough camera.

4.3 TEST ENVIRONMENT

The test environment for the experiment was a windowless room approximately 10ft by 10ft. It was important for the room to be windowless to standardize the lighting conditions across all participants. It was found in preliminary testing that changes in lighting conditions affected a person's ability to detect inconsistencies between the real and virtual environments. The room had a table placed against one wall with a white foam board taped to the top of it. This foam board served as a plain background for the black stimuli as well as a quick way to place the round balls in their appropriate positions without them rolling away. There were notches cut into the board to allow the balls to rest in the same position after removing and replacing them. There was also a spot marked on the floor approximately 1m away from the center ball's position on the foam board.

4.4 PARTICIPANTS

A group of 13 participants (5 females, 8 males, mean age: 21.6) were recruited from the student population of the University of Mississippi. 12 of the 13 had either 20/20 uncorrected vision or used glasses/contacts to attain a 20/20 corrected vision. Informed consent was acquired from all participants prior to participation.

4.5 EQUIPMENT

The experiment was conducted using the Oculus Quest 2, shown in figure 5, running the latest firmware. The Oculus Quest 2 utilizes an 1832x1920 resolution per eye LCD refreshing at, in this experiment, 90Hz. The tracking for the headset is provided by an optical inside-out tracking solution. The software for the experiment was developed using Unity 3D version 2022.1.0b5 on a Windows 11 desktop. The physical balls used as stimuli were Python Black Racquetballs.

Figure 5

This figure shows the Quest 2 HMD and the pair of Oculus Touch Controllers.

4.6 ETHICS STATEMENT

All of the methods and procedures employed in the experiment were reviewed and approved by the University of Mississippi Institutional Review Board. It was classified as exempt status which required the following of COVID-19 protocols and the obtaining of informed consent through an IRB approved consent form. Each participant was required to read this consent form which included the purpose of the study, what would be expected from the participant, time required, risks involved, notice of right to withdraw, and a confidentiality statement. All information was anonymized at the time of collection.

4.7 METHODS

The study required approximately 30 minutes per participant. Before entering the experiment room, each participant was required to use hand sanitizer and verbally affirm that they had not been exposed to anyone nor were they experiencing any symptoms related to COVID-19. After this affirmation, each participant was asked to read and sign an informed consent waiver to participate in the experiment. Once informed consent had been obtained, the participant was asked to complete a Kennedy-Lane Simulator Sickness Questionnaire [12]. After the questionnaire, the participant was shown the buttons on the Oculus Touch Controllers that would be used in the experiment.

The step-by-step process of the experiment and what would be required from the participant was then reviewed. The participant was then told exactly what to expect from the upcoming trials. This involved teaching a person what is meant by "real" objects and "virtual" objects. This can be tricky to understand when not familiar with passthrough camera views in VR headsets. The next step was to inform the participant of the Two Alternative Forced Choice question they would need to answer: "Were all of the objects on the table in the previous trial real?" Once everything had been explained, the participant was asked if they had any questions and that they fully understood the task they were to complete. The final step before beginning the experiment was to ask the participants to repeat the question back that they were to answer as well as point out the buttons that would be used in the experiment.

The participant was then told to begin whenever they were ready. Once the participant began, they would see the passthrough stimulus with either all objects being real or one object being replaced with a virtual object. The participant would then answer the question of whether the objects were all real or not and then wait while the principal investigator changed the layout of the physical balls according to a predetermined, pseudorandom order. The participant would wait until the principal investigator informed them to continue. This was repeated for all 18 trials of the experiment.

After the experiment was complete, the participant was immediately asked to complete the post-experiment follow-up of the Kennedy-Lane Simulator Sickness Survey. Once the responses had been collected, the participant was asked to fill out a general demographics form containing questions about vision and previous experience with virtual reality headsets.

The data being collected during the experiment was: coordinates of the participant's head, coordinates of the virtual ball, target eccentricity, response (real or not), and response time. This data as well as the data from the simulator sickness survey and demographics survey were anonymized upon creation with the only linking factor being a participant number that was randomly given to each participant.

After the experiment was finished, the headset and controllers were thoroughly cleaned with cleaning wipes to avoid transmission of illness between participants. This was done in accordance with the University of Mississippi Covid-19 protocols, but also because it has been found that head mounted displays are considerable vectors for the transmission of bacteria when not properly sanitized between users [13].

CHAPTER 5: RESULTS

5.1 DIMINISHED REALITY PERCEPTION

Figure 6 and 7 show the likelihood of a participant responding real by each position. An analysis of variance (ANOVA) shows that none of conditions significantly differ: $F(5, 60)=1.391$, $p=0.241$. The results don't change if we look at Unsigned Eccentricity: $F(3, 36)=1.420$, $p=0.253$. The data shows that no condition significantly differed from chance except the condition where all items were truly real. This is consistent with typical findings in search literature; participants were more accurate in target absent trials. The data showed that when all objects were real, the participants' responded, on average, that the scene was all real 69% of the time.

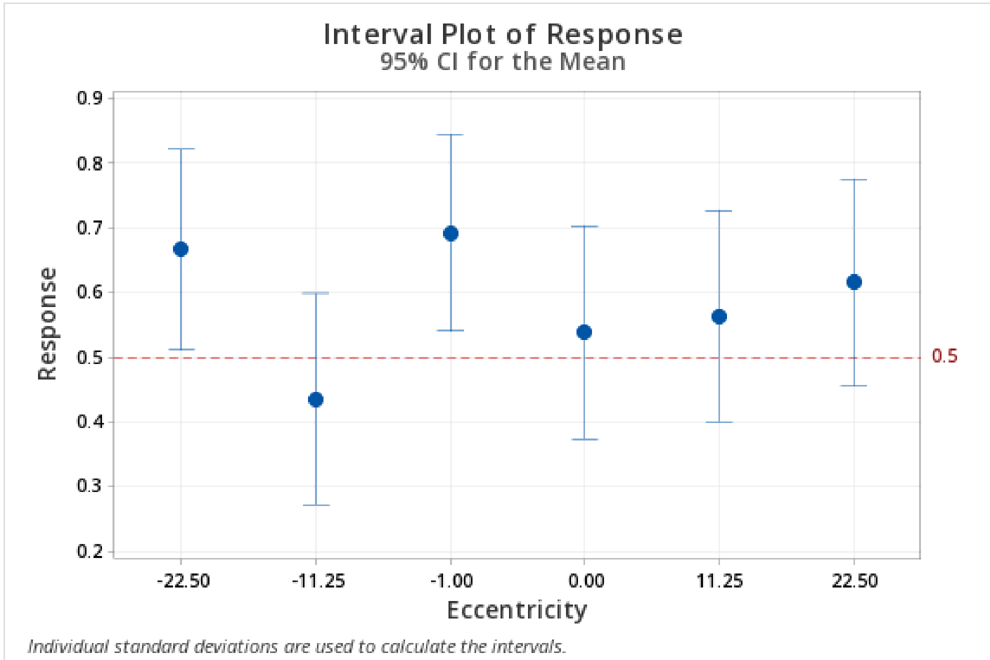

Figure 6

This plot shows the mean of response 1 or 0 for each eccentricity as well as the target-not-present condition (-1).

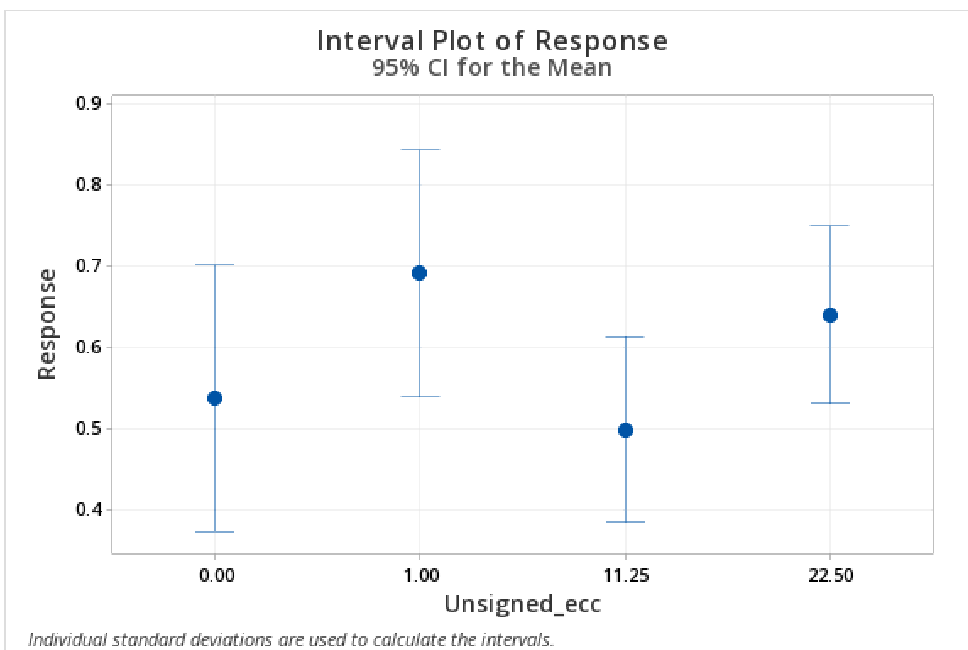

Figure 7

This plot is the same as Figure 6, but it shows the median response for unsigned eccentricity. Here 1 is the target-not-present condition.

5.2 RESPONSE TIME AS A FUNCTION OF TRIAL PROGRESSION

Figure 8 shows a downward trending curve of response time decreasing as the participant progressed through trials. We hypothesize that the initial response times are high while the participant was becoming familiar with the task. Interestingly, the median response time from all participants seems to have converged to a similar level around 2 seconds after 4 trails had been completed. For all of the remaining trials after this convergence, the median response time remained at this level.

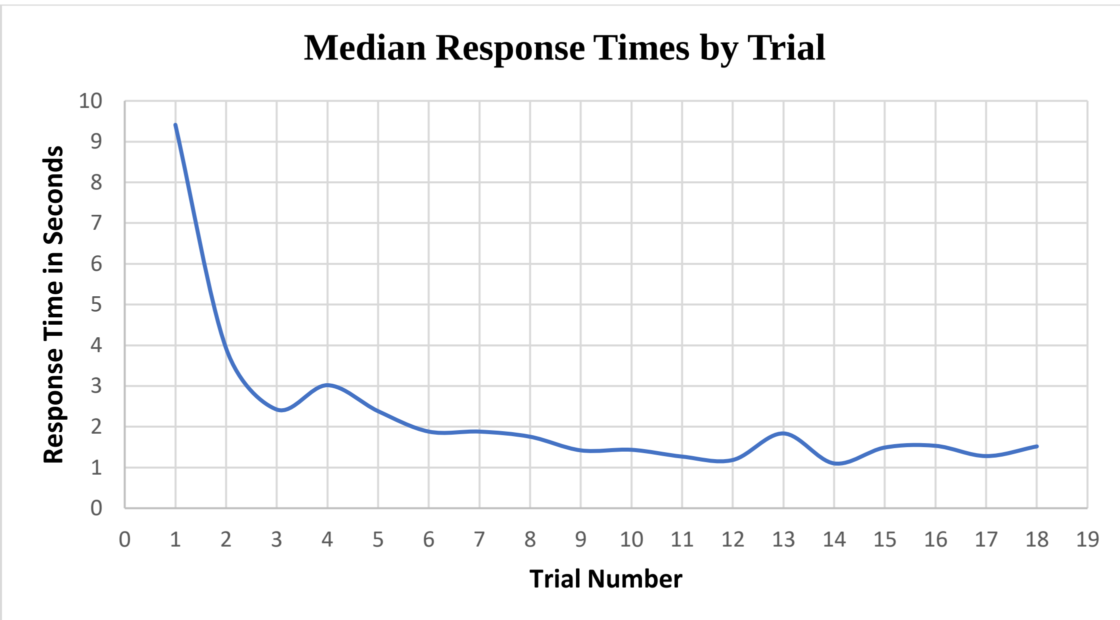

Figure 8

This plot shows the median response of each participant as they progress through the 18 trials of the experiment.

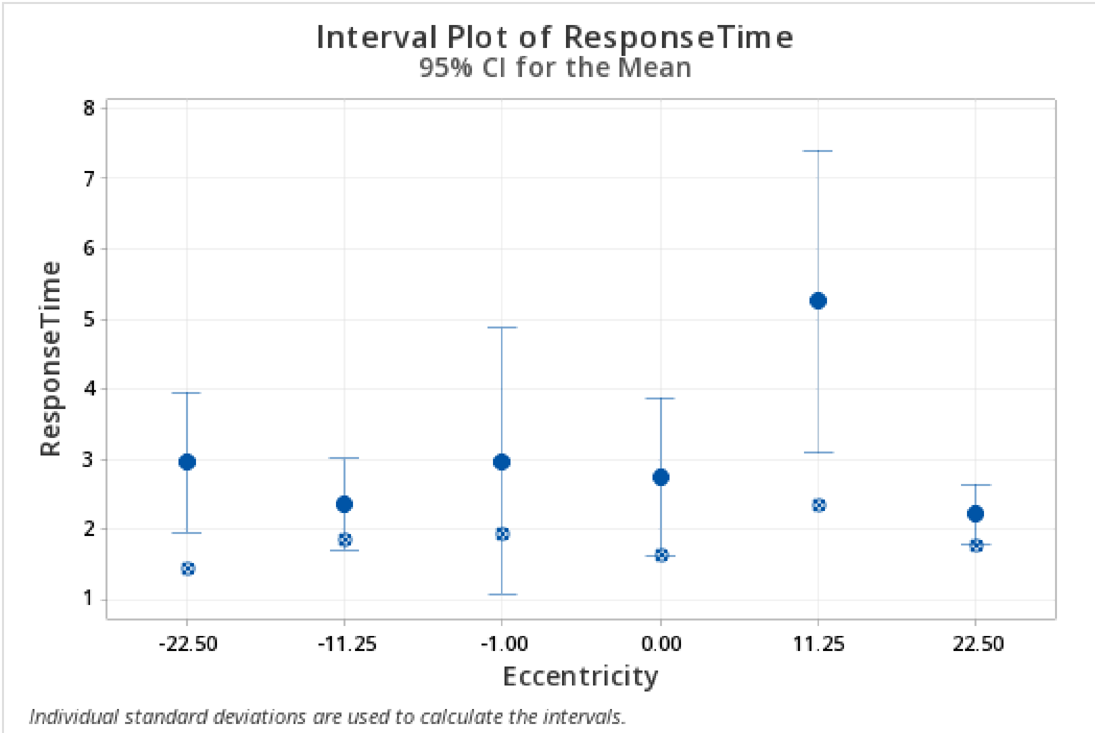

Figure 9

This plot shows the mean (completely shaded) and the median (semi-shaded) response times for each eccentricity. Here, the -1 again represents the target-not-present condition.

5.3 SIMULATOR SICKNESS

A slight increase in general discomfort was observed in 4 of the 13 participants. This is accounts for approximately 31 percent of the sample population experiencing slight general discomfort after running through the 18 trails of the experiment. One participant (approx. 8%) reported slightly increased fatigue after the experiment, and 3 participants (approx. 23%) reported a decrease from slightly fatigued to no fatigue. This is interesting because the experiment did not last that long, and required intense focus to attempt to search for and recognize the camouflaged virtual object among several real objects. New slight headaches were reported by 2 participants (approx. 15 percent), and 2 participants (approx. 15%) reported existing slight headaches had disappeared after completing the experiment. 4 participants (approx. 31%) reported slight increases in eye strain, and, surprisingly, one participant reported a slight decrease in eye strain from the experiment. There was no reported increase in salivation; however, there were 3 participants (approx. 23%) who reported decreases from slightly increased salivation to normal salivation after the experience was complete. 3 participants (approx. 23%) reported increased sweating from the experience. There were 2 participants (approx. 15%) who reported slightly increased nausea from a baseline of none to slight. 2 participants (approx. 15%) observed an increase in their difficulty concentrating, while 3 participants (approx. 23%) observed decreases in difficulty concentrating. The participants who had increases only increased slightly. Two of the participants who experienced decreased difficulty concentrating improved from a slight difficulty to none, but there was one participant who reported a large decrease from moderate difficulty concentrating to no difficulty. This aligns with what was observed in the fatigue category. The required focus and relatively short duration of the experience seems to have caused participants to become more alert. Fullness of head was reported to have increased in 3 participants (approx. 23%), but 3 participants also reported a decrease in feeling of fullness of head. Slightly blurred vision compared to no blur was reported by only one participant, and a decrease from slightly blurred to no blur was observed in one participant. Dizziness with eyes open was only observed in one participant. This is most likely due to the aforementioned lack of head movement with a static fixation point. Dizziness with eyes closed however was shown to have increased in 3 participants. Two participants reported slight increases in vertigo, but one participant reported an increase from none to moderate vertigo. Stomach awareness was only reported in one participant who increased from none to moderate. This participant also reported dizziness closed eyes and vertigo. They seem to have been more sensitive to the effects of SS. There were no participants who reported burping as a side effect of using the HMD.

5.4 BIASES

There was a clear bias toward answering that all of the objects were real. This was the response that was correlated with the right-hand trigger. While this is most likely not causal, all of the participants provided the information that they were right-handed on the demographics survey.

Two participants marked that every trial was composed of entirely real objects even though only 3 out of the 18 trials were truly all real. One of these participants responded to the demographics survey that they were colorblind and had a hard time distinguishing similar shade. The other participant marked that they had very bad uncorrected vision, but were wearing contacts.

CHAPTER 6: DISCUSSION

6.1 ECCENTRICITY

The results of the study show that target eccentricity did not contribute to a statistically significant difference in participants' accuracy when identifying a trial as all real or not. It can be assumed that 300ms of stimulus exposure was not enough time for participants to complete this specific visual-search task where targets are only slightly distinguishable from non-targets as well as the background. This is interesting because previous work has shown that in cognitively demanding visual-search experiments, such as a scene being flashed for just 20ms, visual processing can occur in under 150ms [11]. This implies that the task of discerning whether a virtual object was present in the scene was quite difficult and required high amounts of cognition.

Several participants made comments about the experiment after it was over. Several said they couldn't tell a difference between the trials at all. This could be because the diminished reality shader applied to the virtual stimulus was effective in mimicking the appearance of the real stimuli, but further research would need to be conducted before accepting this as the correct assumption. It is far more likely that the exposure time needed to be extended. This more likely conclusion is supported by the fact that nearly all participants expressed concern that the exposure time was too short which resulted in them being forced to make a guess.

6.2 RESPONSE TIME

Response time is typically correlated with cognitive load; the more difficult a task is, the longer it will typically take for participants to respond [17, 18]. The results of this experiment show that response time generally began very high in the first 3-4 trials participants were exposed to. This can be thought of as the learning period; it took a few trials for participants to become familiar with the task and get comfortable answering the question. The interesting thing is that participant response time converged to around 1.5 seconds after around 6 initial trials. This means that after one full cycle of the 6 different stimuli, the median response time for the following two repetitions were consistent. This aligns with previous research which found that performance in visual-search tasks could be improved after some initial training [14, 15, 16]. This one and a half second response time is a possible indicator that the visual-search task was quite difficult for participants. The stimulus was shown for only 300ms, but it typically took 1510ms to 2000ms to respond after the scene had disappeared. So, using response time as an indicator of cognitive load, the task must have been quite difficult. The question of why the task was so cognitively challenging naturally follows.

What can be presumed to be the causes of this difficulty can be condensed into three main categories: visual fidelity, color, and exposure time. Visual fidelity encompasses several different factors such as the lack of defined shapes. The passthrough cameras on the headset tend to blur the edges of real-world objects making them less salient from their background. The scene also contains unintentional and intentional distractors. While not intentionally incorporated as a distractor, it can be assumed that the visual noise or graininess included in the video feed by the headset's cameras served to create a more difficult environment for participants to discern distinct objects. The next factor that contributed to the difficulty of the scene is the lack of color. As previously mentioned, the cameras on the headset capture video in grayscale. This subspace of color allows for less distinctness and difference among both the stimuli and the background in the scene. The brain cannot use color cues besides differences in hue. This is not something that the brain can quickly distinguish, and it makes the task especially difficult for those who have ocular deficiencies that impede ability to distinguish similar colors. The final part of the experiment that caused the task to be difficult was the exposure time. The time each scene was exposed to the participant was 300ms. This amount of time should have been enough for the participants to complete a simple visual-search task, but upon evaluation of the results, it seems it was not long enough for such a difficult task. All of these factors work together to create a scene which is visually cohesive throughout with no easy way for the brain to quickly and consistently differentiate whether objects are real or superimposed into the camera feed.

6.3 SIMULATOR SICKNESS

"Simulator sickness (SS) in high-fidelity visual simulators is a byproduct of modem simulation technology. Although it involves symptoms similar to those of motion-induced sickness (MS), SS tends to be less severe, to be of lower incidence, and to originate from elements of visual display and visuo-vestibular interaction atypical of conditions that induce MS." [12] Before and after each participant wore the HMD and stepped through the entire experiment, a Kennedy-Lane Simulator Sickness Questionnaire was administered. The questionnaire is given twice; the first time is to establish a baseline of each symptom of simulator sickness for each participant, and the second is used to measure changes in each symptom of SS. These symptoms include:

general discomfort, fatigue, headache, eye strain, increased salivation, sweating, nausea, difficulty concentrating, fullness of head, blurred vision, closed eye dizziness, open eye dizziness, vertigo, stomach awareness, and burping.

The general results were that most participants were unchanged by the experiment except for a few notable exceptions. The notable changes were only slight increases or decreases. In no run was a severe adverse reaction observed. This is most likely due to the nature of the experiment. The experiment required no head movement. Head movement is the most common cause of SS. The duration of the experience is also a great indicator of whether SS will occur. This experiment also had a maximum duration of 15 minutes in the headset which can also be attributed to the experiment's minimal impact on participants.

CHAPTER 7: CONCLUSION AND FUTURE WORK

The results of this study show that there was essentially no target eccentricity condition which created a divergence from a random guess from the participant besides the condition where all objects were truly real. These results are very interesting because they indicate that humans are not able to quickly discern virtual, superimposed objects from a passthrough scene. This is quite exciting because it supports the hypothesis of this paper and implies that a similar method of *diminished virtual reality* could be utilized in mixed reality experiences on the Oculus Quest to create a more cohesive and convincing blend of real and virtual worlds. While the eccentricity effect was not present as pronounced effect and the experiment could be improved upon to discern more, it is exciting to see that the task was challenging for participants. This research, while not definitive, helps answer the initial question that prompted this research: “Is there a way to create a better blend of real and virtual worlds in a mixed reality experience.”

A follow-up to this study would need to evaluate the exposure time needed to detect the target object. In this follow-up, Two Alternative Forced Choice would be utilized again, but using exposure time as the independent variable instead of eccentricity. Another thing that would be worth changing is requiring participants to complete a practice trial to familiarize themselves with the task. This would, most likely, result in clearer trends emerging.

BIBLIOGRAPHY

[1] M. Carrasco et al, "The eccentricity effect: target eccentricity affects performance on conjunction searches," Perception & Psychophysics, vol. 57, (8), pp. 1241-1261, 1995.

[2] T. Alsop, “Steam user VR headset share by device 2022,” Statista, 14-Feb-2022, Available: https://www.statista.com/statistics/265018/proportion-of-directx-versions-on-the-platform-steam/.

[3] S. Mori, S. Ikeda and H. Saito, "A survey of diminished reality: Techniques for visually concealing, eliminating, and seeing through real objects," IPSJ Transactions on Computer Vision and Applications, vol. 9, (1), pp. 1-14, 2017.

[4] Oculus, “What is Passthrough?”, Available: https://support.oculus.com/articles/in-vr-experiences/oculus-features/what-is-passthrough/.

[5] Mann S (1994) Mediated reality. M.I.T. Media Lab Perceptual Computing Section, Cambridge, Ma. Technical Report TR 260

[6] Wolfe, J.M., O’Neill, P., & Bennet, S.C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. Perception and Psychophysics, 60(1), 140-156.

[7] J. Duncan and G. W. Humphreys, "Visual Search and Stimulus Similarity," Psychological Review, vol. 96, (3), pp. 433-458, 1989.

[8] A. S. Hess et al, "On the hunt: Searching for poorly defined camouflaged targets," PloS One, vol. 11, (3), pp. e0152502-e0152502, 2016.

[9] M. B. Neider and G. J. Zelinsky, "Searching for camouflaged targets: Effects of target-background similarity on visual search," Vision Research (Oxford), vol. 46, (14), pp. 2217-2235, 2006.

[10] M. Carrasco et al, "Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency," Vision Research (Oxford), vol. 38, (3), pp. 347-374, 1998.

[11] S. Thorpe, D. Fize and C. Marlot, "Speed of processing in the human visual system," Nature (London), vol. 381, (6582), pp. 520-522, 1996.

[12] R. S. Kennedy et al, "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness," The International Journal of Aviation Psychology, vol. 3, (3), pp. 203-220, 1993.

[13] Benjamin Creel, Caitlin J. Rinz-Jones, Adam Jones, and Colin Jackson. 2020. Bacterial Load of Virtual Reality Headsets. In 26th ACM Symposium on Virtual Reality Software and Technology (VRST '20). Association for Computing Machinery, New York, NY, USA, Article 34, 1–8. DOI:https://doi.org/10.1145/3385956.3418958

[14] Boot W.R., Neider M.B., & Kramer A.F. (2009). Training and transfer in search for camouflaged real-world targets. Attention, Perception & Psychophysics, 71, 950–963.

[15] Chen X. and Hegdé J. (2012). Learning to break camouflage by learning the background. Psychological Science 23:1395. pmid:23064405

[16] Neider M.B., Boot W.R., & Kramer A.F. (2010). Visual search for real-world targets under conditions of extreme target-background similarity: Exploring training and transfer in older adults, Acta Psychologica, 134, 29–39. pmid:20038458

[17] E. Cooper-Martin. Measures of cognitive effort. Marketing Letters, 5(1):43–56, 1994.

[18] Finney, H. and Jones, J.A., 2020, March. Asymmetric effects of the Ebbinghaus illusion on depth judgments. In 2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR) (pp. 573-578). IEEE.