A Comparison of Eye-Head Coordination Between Virtual and Physical Realities

Kevin Pfeil
University of Central Florida
Orlando, Florida
kevin.pfeil@knights.ucf.edu

Eugene M. Taranta II
University of Central Florida
Orlando, Florida
etaranta@gmail.com

Arun Kulshreshth
University of Louisiana at Lafayette
Lafayette, Louisiana
arunkul@louisiana.edu

Pamela Wisniewski
University of Central Florida
Orlando, Florida
pamwis@ucf.edu

Joseph J. LaViola, Jr.
University of Central Florida
Orlando, Florida
jjl@cs.ucf.edu

ABSTRACT

Past research has shown that humans exhibit certain eye-head responses to the appearance of visual stimuli, and these natural reactions change during different activities. Our work builds upon these past observations by offering new insight to how humans behave in Virtual Reality (VR) compared to Physical Reality (PR). Using eye- and head- tracking technology, and by conducting a study on two groups of users - participants in VR or PR - we identify how often these natural responses are observed in both environments. We find that users statistically move their heads more often when viewing stimuli in VR than in PR, and VR users also move their heads more in the presence of text. We open a discussion for identifying the HWD factors that cause this difference, as this may not only affect predictive models using eye movements as features, but also VR user experience overall.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; • **Computing methodologies** → *Tracking*;

KEYWORDS

Eye tracking, head tracking, user study

ACM Reference Format:

Kevin Pfeil, Eugene M. Taranta II, Arun Kulshreshth, Pamela Wisniewski, and Joseph J. LaViola, Jr.. 2018. A Comparison of Eye-Head Coordination Between Virtual and Physical Realities. In *SAP '18: ACM Symposium on Applied Perception 2018, August 10–11, 2018, Vancouver, BC, Canada.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3225153.3225157

1 INTRODUCTION

In recent years, we have seen a resurgence of immersive *virtual reality* (VR) technology, with the advent of affordable and available Head-Worn Displays (HWDs), such as the Oculus Rift, HTC Vive, and Samsung GearVR. These devices have been marketed to the common user as sources of entertainment, allowing people to have new and exciting experiences. VR HWDs are constantly being designed with better features, and now we can find iterations that utilize eye tracking technology. It is important to understand how people form eye-head coordination in VR, as this could affect the success of eye-tracking technology for its intended tasks. Similarly, we need to determine how eye-head coordination in physical reality (PR) compares to that in VR. This will help us understand how to simulate certain tasks that garner specific eye responses.

In this paper, we discuss an exploratory experiment designed to invoke natural eye-head coordination for viewing basic stimuli. Participants - in either VR or PR - are asked to perform basic visual attention tasks. In order to measure these behaviors, we record eye and head movements, and then classify them according to patterns previously seen by prior research. Our hypotheses for the experiment are as follows:

- H1: Eye-head coordination differs between VR and PR for the same tasks
- H2: In VR, simple looking invokes different eye-head coordination than alphanumeric character identification
- H3: In PR, wearing a simulated HWD invokes different eye-head coordination than unrestricted gazing

Our results indicate that users exhibit statistically contrasting eye-head responses to similar stimuli in VR and PR; particularly, VR users tend to move their heads towards a stimulus more often than PR users. This is an interesting finding that helps us understand natural human behavior in VR, but it also highlights an underlying phenomenon which may affect VR user experience.

2 RELATED WORK

The eyes have been studied for many decades [Bizzi 1974] [Dodge 1921] [Land 1992] [Mourant and Grimson 1977] [Shackel 1960] [Yarbus 1967]. In recent years, the eyes have been studied for human-computer interaction, including helping the disabled [Heikkilä 2013] [Hornof et al. 2004] [Hornof and Cavender 2005] [Ward et al. 2000] [Ward and MacKay 2002], as additional input for gaming [Isokoski et al. 2009] [Krejtz et al. 2014], understanding gaze patterns in various environments [Djamasbi et al. 2010] [King et al.

SAP '18, August 10–11, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5894-1/18/08…$15.00
https://doi.org/10.1145/3225153.3225157

2017], and general HCI use [Duchowski 1998] [Jacob and Stellmach 2016] [Jacob 1990].

Nearly one century ago, Dodge [Dodge 1921] conducted an experiment to study how the eyes reacted to head movements. The term "compensatory eye movement" was used to describe the phenomenon [Dodge 1921], and Shackel [Shackel 1960] also saw this behavior; this essentially leads into what we today call the Vestibulo-Ocular Reflex (VOR). The VOR is the behavior of our eyes to react to head movements, by moving in a counter direction. This function allows us to see clearly while our head moves; without it, we would be unable to focus on a stimulus when rotating our head. Di Girolamo et al. [Stefano Di Girolamo 2001] observed an effect on the VOR by VR, indicating that natural human behaviors are inhibited after VR usage. This finding could have been an artifact of the used technology - their HWD had a fraction of the screen resolution that can be found on today's hardware - but we are currently unsure if this is the case. Our work is somewhat similar to that of Di Girolamo's, in that we are observing how VR affects natural human eye movements. Ours differs, though, in that we are studying natural behaviors during usage, not after. We also tackle this problem with less equipment; instead of using an electronystagmograph (ENG) device, we utilize eye- and head- trackers to record and classify movements based on prior categorizations.

Existing literature reuses the terms "classical" and "predictive" eye movements to describe types of eye-head coordination [Bizzi 1974] [Mourant and Grimson 1977] [Shackel 1960]. *Classical* movements refer to the common occurrence of when the eyes move before the head towards a stimulus; *Predictive* movements refer to when the head moves before the eyes. Bizzi noted that the head only moved before the eyes when a stimulus was in a position that a subject could anticipate [Bizzi 1974]. Classical and predictive movements were again observed by Mourant [Mourant and Grimson 1977] during an automobile driving task. In this scenario, the predictive movements occurred very often, perhaps because drivers knew exactly where to look - at the side and rear view mirrors - before making eye movements. It also makes sense that drivers want to keep eyes forward for as much time as possible; so naturally, the head would move towards a mirror before the eyes. Shackel's study revealed predictive movements even in an open environment, although they were very rare [Shackel 1960]; so it seems that these can occur even outside of a known and anticipated environment.

In our work, we reuse the classical and predictive classifications in order to help understand the types of eye-head coordination that are used in both VR and PR. To our knowledge, there has not been any other work to help understand these differences in VR. Our pilot studies have revealed the necessity for one more classification - "Eyes Only" - to denote when a user views a stimulus without moving the head. This can technically fall into the "classical" bucket, but it warrants its own unique category, as we have found it to be a common occurrence. This assists us in identifying if our main hypothesis is correct. In the remainder of our paper, we will refer to these three categories, and summarize them here:

- **Eyes Only** (head never moves)
- **Classical** (eyes move before head)
- **Predictive** (head moves before eyes)

As Di Girolamo et al. observed an effect on the VOR after using VR [Stefano Di Girolamo 2001], and since Shackel observed predictive movements while wearing heavy equipment on the head [Shackel 1960], we expect that there exists one or more factors of today's immersive VR technology that affects eye-head coordination when compared to PR. We thus expect this work to be the first of many efforts to identify factors which may cause a difference in natural behaviors.

3 EXPERIMENTAL DESIGN

To test our hypotheses, we conducted a 2x2 mixed-design study. The groups - VR and PR - were between-subjects. Each group had two conditions, which were within-subjects. For VR, we wanted to explore different virtual scenarios and see if eye-head coordination differed between them. The first scenario, "Look", is a simple visual attention task. The second scenario, "Read", is the same as Look, except it has the user read one letter off of the stimulus. Data from this group helps us answer hypothesis H2. The PR group had two scenarios which replicated the VR Look scenario. The first, called "Normal", is essentially a PR adaptation of the VR scenario Look. The second, called "Modified", is the same, except the user wears a shell of an HWD. Data from this group helps us answer hypothesis H3. Thus, we had 2 groups, each with 2 conditions:

- **VR_Look** - Look scenario in VR
- **VR_Read** - Read scenario in VR
- **PR_Normal** - No HWD condition in PR
- **PR_Modified** - Modified HWD condition in PR

A comparison of data between **VR_Look** and **PR_Normal** helps us answer hypothesis H1. For the VR group, the independent variables were the virtual scenarios and the stimulus location (see Figure 1). For the PR group, the independent variables were the HWD type and stimulus location (see Figure 1). For both, the dependent variable is the category of eye-head coordination used to look at the stimuli.

3.1 Virtual Scenarios

For the VR group, we developed two virtual scenarios - **VR_Look** and **VR_Read** - to see if different visual attention tasks could affect eye-head coordination. Overall scenario design was influenced by other research (see Sibert et al. [Sibert and Jacob 2000]) that used a desktop environment. Ours is a modified, 3D VR adaptation of this prior design. The common task in these scenarios is to view these stimuli in a randomly generated order.

For these scenarios we used Unity3D to generate a simple 5x5 grid of spheres equidistant from the user; see Figure 1. The grid was empirically designed such that the grid filled up the HWD's view-port, and all spheres could be seen while looking at the center. As such, users would not need to move their head to complete any tasks. Each sphere was assigned a vertical and horizontal angle, in multiples of 14.4 and 16 degrees, respectively, and then were translated 7 meters away from the camera. In essence, this aligned each stimulus on an invisible hemisphere that would wrap around the camera. The visual angle for each sphere was approximately .142615 radians, or 8.17 degrees. We wanted the visuals to be comfortable to look at through the duration of the experiment, so the sphere color was white, and the background color was a shade of

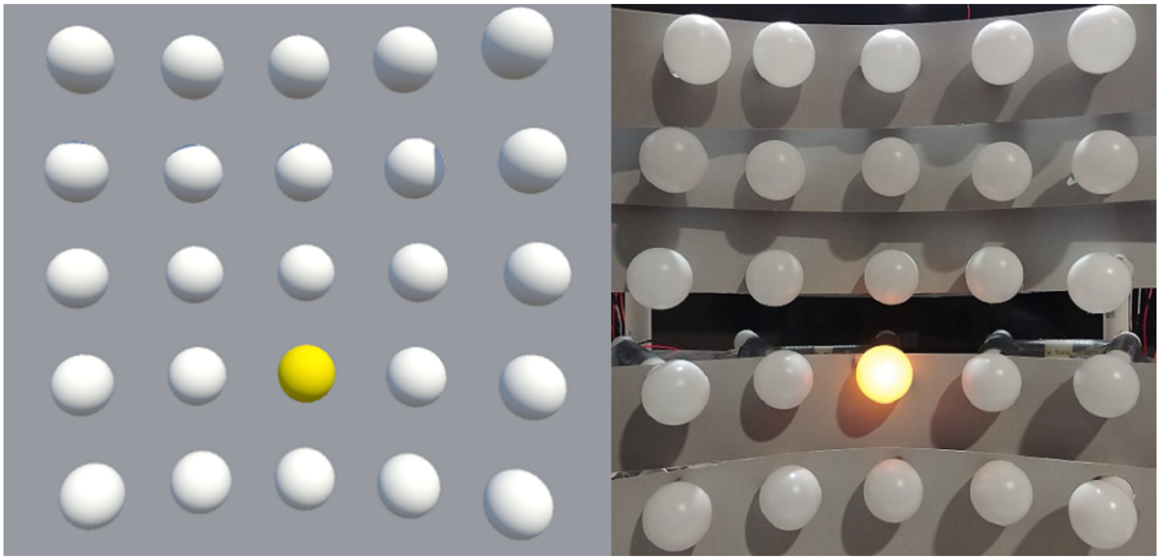

Figure 1: A screenshot of the virtual stimulus layout (left) and a view of the physical stimulus rig (right).

grey. When a sphere was selected as a stimulus, it turned yellow. **VR_Read** was the same as **VR_Look**, except each sphere had a random letter of the English alphabet overlaid on it. These letters were randomized between each measurement. In both scenarios, the user was able to move their head to look around. The center sphere was used to help align the users' head rotation and gaze, and was not used for measurement.

3.2 Physical Reality Stimulus Rig

For the PR group, we constructed a stimulus rig that imitated the **VR_Look** scenario. We used the visual angle of the virtual stimuli - .142615 radians - in order to construct our device. This fixed angle is related to both the size of stimuli as well as user distance from the device. By fixing one value, the other must adjust. We elected to use ping pong balls to represent the spheres and thus measured the resulting distance that the user would need to be from our device. The ping pong balls were 38mm (1.5in) in diameter and that resulted in a user distance of 26.6cm (10.47in). In order to light up our stimuli, we inserted LED lights that wired to an Arduino Uno. The Uno turned on a specific stimulus by communicating with Unity3D over a serial port, via USB.

We used plexiglass and PVC pipes in order to construct shelves on which to place the stimuli. Using the known angles (14.4 degrees vertical and 16 degrees horizontal) and the known 10.47in distance between the user and each stimuli, we constructed the rig as closely as possible to replicate the VR stimuli. To imitate the grey background, we used a simple, thin, gray poster sheet in order to create strips of shroud. The shroud pieces blocked the wiring, PVC, and other irrelevant or potentially distracting objects from the users' view. However, we left a gap between the 3rd and 4th rows, to accommodate our capture devices. See Figure 1 for a sample of the users' viewpoint during the study.

3.3 Subjects

3.3.1 Virtual Reality. 21 participants were recruited from the University of Central Florida student body. Data from one participant was discarded due to hardware failure. Our resulting participant pool consisted of 5 females and 15 males. The average age was 21.3 (median 20). We asked our participants if they felt any head or neck pain before beginning the study, but none did. 11 users reported having some degree of background using a HWD such as the Oculus Rift, HTC Vive, or even Google Cardboard. The remaining 9 did not have any experience. 10 users reported being near-sighted, 1 was far-sighted, and the remaining 9 considered themselves as having normal eyesight.

3.3.2 Physical Reality. 22 participants were recruited from the University of Central Florida student body. Data from two participants was discarded due to hardware failure. Our resulting participant pool consisted of 7 females and 13 males. The average age was 23.5 (median 24). We asked our participants if they felt any head or neck pain before beginning the study, but none did. 8 users reported being near-sighted, 1 was far-sighted, and the remaining 11 considered themselves as having normal eyesight.

3.4 Apparatus

3.4.1 Virtual Reality. The Unity3D game engine version 2017.1.0f3 was used to develop and display VR scenarios. They were run on a laptop computer with a core i7 processor at 2.6GHz, with 12GB RAM, and an NVIDIA GeForce GTX 970M graphics card. Our HWD used in these scenarios was an HTC Vive that had an embedded eye tracker, per Tobii Technologies[1]. The HWD weighed approximately 1.3lbs (580g). The eye tracking module was binocular at 120Hz, and head tracking occurred at 90Hz. We were able to record user eye gaze and head rotation, per frame, through the Unity3D engine. We

[1]https://www.tobiipro.com/vr/

were also able to consistently achieve a rate of 75 frames-per-second, or roughly one frame per 13ms.

3.4.2 Physical Reality. The Unity3D game engine was again used to develop a data recording module which interfaced with our custom stimulus rig, described previously. This was run on a desktop computer with a core i7 processor at 3.4GHz, with 16GB RAM.

We took apart an Oculus Rift DK1 by removing the circuitry and screen; all that remained were the casing and head bands. We used this as our modified HWD. The disassembled HWD weight was approximately .33lbs (148g). We used the Polhemus Patriot in order to track the head rotations of our users. We sewed one of the sensors into the band of a lightweight hat. The total weight of the hat and sensor was approximately .27lbs (122g). Head tracking occurred at 60Hz.

We originally chose the Tobii EyeX to perform eye tracking, but we found that it was unable to consistently track our users' eyes. We thus elected to use a small video camera (360Fly), to simultaneously record the users' eyes and debugging information displayed on the computer monitor, which informed us of the active stimulus. The camera recorded at 30FPS. The camera was placed on a mount behind the stimulus rig and was visible between the rows 3 and 4 in the stimulus rig.

3.5 Experiment Procedure

Our participants were all given the same set of instructions by the experimenter. First, an overall description of the study was communicated. Next, a short demographics survey was administered. Then the participant was assigned an order of conditions in a counter-balanced design. For the VR group, users were seated in a standard location within the Vive's tracking space so to keep the environment constant. For the PR group, users were seated in front of the stimulus rig. Users were asked to adjust their seat if needed, so that "looking straight forward" resulted in looking at the center sphere or ball.

Directions for completing the study were to simply view the stimuli in any manner the user saw fit, and inform the experimenter when they have. For **VR_Read**, the participants were asked to read aloud the random letter of the highlighted sphere. After acknowledging the instructions, the participant donned the appropriate headgear. For the VR group, users were sat in a specific spot in the HTC Vive interaction space, for consistency. Likewise, for the PR group, the users were sat in a specific spot in front of the rig, and the distance between the head and the rig was measured, to verify consistency. For both, if the user began to slouch, lean, or otherwise move from the standard spot, we politely asked them to reset their posture.

The following steps were then completed for all conditions: a random stimulus was turned on, by becoming highlighted, and for **VR_Read**, all letters in the scene were again randomized. The user then viewed the stimulus and verbally indicated that the task was complete. The experimenter then pressed a keyboard key, which resulted in the center stimulus becoming highlighted, to bring participant gaze back to the normal position. When the participant was again looking at this center stimulus, the experimenter pressed a keyboard key, to move to the next random stimulus. This process continued until every stimulus was used.

After a condition was completed, the user was allowed to remove any headgear and take a short break, if needed. The experimenter then prepared the next scenario and repeated the previous steps. After a participant completed both conditions in their group, a short questionnaire was given. In particular, we wanted to capture any discomfort the task may have caused. We asked the users if they experienced any eye strain, neck pain, dizziness, or nausea, each on a 7-point Likert scale. The length of the study per participant was approximately 15 minutes. We compensated our participants with $5 USD in cash. After each user, the headgear was sanitized using rubbing alcohol.

3.6 Data Classification

After collecting the data, we needed to perform classification. Due to the different technologies used to capture VR and PR data, the analysis was done differently between environments. However, we enforced standards for tagging based on the types of eye-head movement previously described:

- If the head never moves, then **Eyes Only**; else,
- If the eyes move before the head, then **Classical**; else,
- If the head moves before the eyes, then **Predictive**

We acknowledged that our tracking devices are not without jitter, and we also note that our heads can be slightly moved by simple acts such as breathing and speaking. Using observational data from our pilot studies, we therefore set a 2 degree threshold for head movement. If the head did not move more than 2 degrees, we considered this to be an Eyes Only response. For the others, all that remained was to see what moved first - the eyes or the head. For VR, we converted the eye-head movements into time series plots, and for PR, we watched the recorded videos. For both, we performed an objective check to see which moved towards the stimulus first.

We collected and classified 480 data points for each condition, for a total of 1920 samples (4 x 480). We were unable to classify 9 total samples (3 in VR, 6 in PR) due to difficulty in reading the data.

4 RESULTS

After data collection and tagging, we analyzed the frequency counts of each classification. Using these counts, we elected to use contingency table analysis.

4.1 Virtual Reality vs Physical Reality

For **VR_Look**, 16.3% of the responses were Eyes Only, 81% Classical, and 2.5% movements were Predictive. For **VR_Read**, 7.9% of the responses were Eyes Only, 90% were Classical, and 1.5% were Predictive. See Figure 2 for a graphical depiction.

For **PR_Normal**, 64% of the responses were Eyes Only, and the remaining 36% of the responses were Classical. For **PR_Modified**, 60% of the responses were Eyes Only, and the remaining 40% were Classical. We did not observe any predictive movements during any measurements. See Figure 3 for an illustration.

To answer hypothesis H1, we used data from **VR_Look** and **PR_Normal**, and performed an omnibus χ^2 test on the frequencies of occurrences for each category. We found a significant difference overall ($\chi^2(2, N = 953) = 231.485, p < .0001$). We then performed a

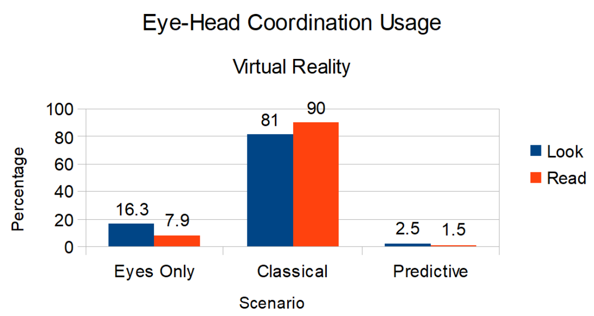

Figure 2: Percentage distribution of eye-head coordination in VR. The majority of movements were Classical.

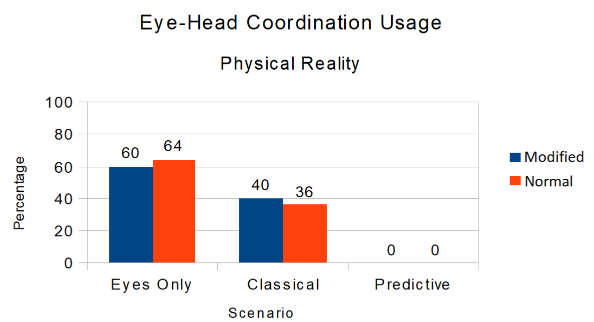

Figure 3: Percentage distribution of eye-head coordination in PR. The majority of movements were Eyes Only.

post-hoc analysis between Eyes Only vs Classical, and we still found a significant difference ($\chi^2(1, N = 941) = 219.465, p < .0001$). We did not perform post-hoc tests involving Predictive, as that response was rare.

We also wanted to examine the **VR_Look** and **PR_Modified** conditions; an omnibus χ^2 test revealed a significant difference here as well ($\chi^2(2, N = 959) = 199.288, p < .0001$). Post-hoc analysis between Eyes Only and Classical again revealed a significant difference ($\chi^2(1, N = 947) = 187.146, p < .0001$). We did not perform post-hoc tests involving Predictive, as that response was rare.

We thought that one of the reasons for our results is that there are two rings formed around the center sphere (see Figure 1). We would expect that stimuli further away from the center would warrant more head movements. For the outer ring, a χ^2 test reveals that there is significant difference in responses between **VR_Look** and **PR_Normal** ($\chi^2(2, N = 636) = 158.767, p < .0001$). Post-hoc analysis between Eyes Only and Classical reveal a difference ($\chi^2(1, N = 624) = 146.707, p < .0001$). We did not perform post-hoc tests involving Predictive, as that response was rare. For the inner ring, a χ^2 test again reveals a significant difference in responses between **VR_Look** and **PR_Normal** ($\chi^2(2, N = 317) = 77.756, p < .0001$). No post-hoc analysis was needed because there were no Predictive responses in the omnibus test.

4.2 Look vs Read

To answer hypothesis H2, we used data from **VR_Look** and **VR_Read**, and performed an omnibus χ^2 test on the frequencies of occurrences for each category. We found a significant difference overall ($\chi^2(2, N = 957) = 17.463, p < .001$). We then performed post-hoc analyses between Eyes Only vs Classical, and found a significant difference ($\chi^2(1, N = 938) = 16.132, p < .001$). We did not perform post-hoc tests involving Predictive, as that response was rare.

4.3 Normal vs Modified

To answer hypothesis H3, we used data from **PR_Normal** and **PR_Modified**, and performed an omnibus χ^2 test on the frequencies of occurrences for each category. We did not find a significant difference overall ($\chi^2(1, N = 954) = 1.732, p = .188$). No post-hoc analysis was needed because there were no Predictive responses in the omnibus test.

4.4 Post Questionnaires

In VR, we found that the users did not express more than mild discomfort for their neck (Mean = 1.2, SD = 0.62) or eyes (Mean = 2.4, SD = 1.57), and only expressed mild discomfort for nausea (Mean = 1.3, SD = 0.66) and dizziness (Mean = 1.6, SD = 1.31).

In PR, we found that the users did not express more than mild discomfort for their neck (Mean = 1.7, SD = 1.26) or eyes (Mean = 2.5, SD = 1.57), and only expressed mild discomfort for nausea (Mean = 1.3, SD = 0.72) and dizziness (Mean = 1.4, SD = 0.75).

5 DISCUSSION

5.1 Eye-Head Coordination Differs Between Realities (H1)

Our results indicate that there is a difference in the natural eye-head coordination between VR and PR. It is obvious that our VR users tend to rotate their head in some capacity for the majority of tasks, whereas our PR users tend to keep their heads still. These results support hypothesis H1, in that eye-head coordination varies between these environments, at least for simple visual attention tasks within the user's field of view.

Now we must ask - *why* are we finding this difference? Currently, we believe HWD field of view, screen resolution, and weight are the most likely factors that contribute to this difference. While our study only used one VR HWD (HTC Vive), we can still compare the technology used by Di Girolamo et al; their device, the Virtual I/O iGlasses!, "consist of a head piece with two 7-in. full-color LCDs, each having a field of view of 30° .... Each LCD panel has a resolution of 180,000 pixels .... the entire unit weights 450g" [Stefano Di Girolamo 2001]. To compare, the HTC Vive has a field of view of 110°, a resolution of 1080x1200 per eye, and weighs approximately 580g [viv [n. d.]]. Is the effect on the VOR perhaps reduced with newer technologies? These features are drastically different, so that could be the case. It will be interesting to see how future devices affect eye-head coordination in VR environments.

We must also ask one more question - *does it matter* that we have this difference? Eye tracking is already being incorporated into a variety of VR applications, and the results seem favorable. For instance, Gandrud et al. were able to predict with some certainty VR user's path in VR by observing eye-head coordination [Gandrud and Interrante 2016]. Comparatively, Karaman et al. use eye tracking for tablet devices, outside VR, to help recognize what operations users wanted to perform [Karaman and Sezgin 2018]. The commonality between these works is that the eye movements are being used to generate a predictive model - but our results imply that it may not be possible to port over the same predictive model from PR to VR. Further, it may not be possible to port the same model between different VR HWDs, due to their varying features. More work needs to be performed in order to determine if this is the case; if so, it would be absolutely necessary to identify the cause for this difference, and work towards a set of standards that would prevent a drop-off in performance.

Our results also cause concern from an ergonomic standpoint. Prior work has shown evidence suggesting that head rotations, with the additional weight of an HWD, can cause pain, strain, or discomfort on the neck [Knight and Baber 2004] [Wille et al. 2014]. Although our experiment's environment was static and did not fall outside the users' view, we still found our users moving their head more in VR than in PR. Finding the cause for this phenomenon can help prevent unnecessary strain, but we must stress weight reduction for HWD designers, to help prevent user discomfort.

5.2 Different VR Scenarios Affect Eye-Head Coordination (H2)

We found a significant difference in eye-head coordination between our two virtual scenarios, and see that, during the **VR_Read** condition, our participants were more likely to move their heads towards the target when compared to the basic **VR_Look** condition. These results support our hypothesis (H2). None of our participants mentioned that the text was blurry or unreadable, so we believe this could be an indication of natural responses. This result is also interesting as it shows that VR users tend to move their heads towards a simple letter of text; our study cannot answer to full words or sentences, but it would follow that user heads should still move in these cases, and perhaps even more. Again, we invoke the works of Wille et al. [Wille et al. 2014] and Knight et al. [Knight and Baber 2004]; to remove physical burdens from a user while reading, careful text placement should be considered by VR designers, as appropriate. For instance, floating texts (e.g. menus) can be placed in the center of the users' field of view.

5.3 Wearing a Simulated HWD Does Not Affect Eye-Head Coordination (H3)

Contrary to the VR conditions, we did not find a significant difference between PR conditions. The distribution of eye-head coordination for **PR_Normal** and **PR_Modified** were fairly similar, so our hypothesis is not supported. As such, we did not identify a behavioral change occurring when this pseudo-mask was worn. This result does compound with our H1 findings, however; there is at least one factor or property of HWDs that cause eye-head coordination to change between VR and PR, but simply donning a lightweight, generic, head-worn device does not seem to cause this change. This leads us to believe that the most likely contributing factors are screen resolution, field of view, and weight; but it may also be true that this difference stems from stimulus type (virtual vs. real). Our future work will assess these various factors and help to identify the root cause for difference.

6 LIMITATIONS

While we are pleased with our results, we acknowledge that our work has some limitations. Due to the nature of our study, we were unable to use the same data collection devices between environments. This is currently unavoidable, and it is possible that this brought some discrepancy. Data classification was thus separate as well, though by utilizing head tracking, we were able to use the same standards for tagging in both environments. Because of this, we are very confident in all of our tagging for Eyes Only movements, which is enough for us to find statistical difference between that and any other movement.

We acknowledge that this study leaves unanswered questions which can be addressed by future studies. For instance, we do not report the types or magnitudes of head rotations. Due to our different devices, we were not able to truly study any natural gaze patterns other than basic eye-head coordination. Additionally, it is possible that other variables not controlled here may contribute to our findings. For instance, while we controlled visual angle, we did not account for factors such as vergence and disparity. We anticipate future researchers having the ability to perform more in-depth studies; our work achieves the milestone of identifying the problem, but more work is needed to provide a better explanation into the cause and solution.

7 CONCLUSION AND FUTURE WORK

We have presented a statistical comparison of eye-head coordination between VR and PR, and our data provides evidence that there is a difference in the formation of natural responses for these two environments. While we are currently unable to identify the specific cause, we do plan on investigating possible reasons. HWDs have many properties to consider, including weight, screen resolution, and field of view. There are perhaps psychological constructs that cause this discrepancy - for instance, the appearance of physical vs. virtual stimuli may play a role, or perhaps making an abrupt transition from PR to VR affects us. We plan on performing future studies that will help us understand how each of these factors affect natural eye-head coordination. This is pertinent to understanding how we can translate eye tracking models in VR, as well as how we can improve overall user experience.

ACKNOWLEDGMENTS

This work is supported in part by NSF under Award IIS-1638060, Lockheed Martin, Office of Naval Research under Award ONR-BAA15001, and Army RDECOM under Award W911QX13C0052.

REFERENCES

[n. d.]. HTC Vive VR System. https://www.vive.com/us/product/vive-virtual-reality-system. ([n. d.]). Accessed: 2018-01-23.

Emilio Bizzi. 1974. The coordination of eye-head movements. *Scientific American* (1974).

Soussan Djamasbi, Marisa Siegel, and Tom Tullis. 2010. Generation Y, web design, and eye tracking. *International Journal of Human-Computer Studies* 68, 5 (2010), 307 – 323. https://doi.org/10.1016/j.ijhcs.2009.12.006

Raymond Dodge. 1921. The Latent Time of Compensatory Eye-movements. *Journal of Experimental Psychology* 4, 4 (1921), 247.

Andrew T. Duchowski. 1998. Incorporating the viewer's point of regard (POR) in gaze-contingent virtual environments. (1998), 332-343 pages. https://doi.org/10.1117/12.307180

Jonathan Gandrud and Victoria Interrante. 2016. Predicting destination using head orientation and gaze direction during locomotion in vr. In *Proceedings of the ACM Symposium on Applied Perception*. ACM, 31–38.

Henna Heikkilä. 2013. EyeSketch: A Drawing Application for Gaze Control. In *Proceedings of the 2013 Conference on Eye Tracking South Africa (ETSA '13)*. ACM, New York, NY, USA, 71–74. https://doi.org/10.1145/2509315.2509332

Anthony Hornof, Anna Cavender, and Rob Hoselton. 2004. Eyedraw: A System for Drawing Pictures with Eye Movements. In *Proceedings of the 6th International ACM SIGACCESS Conference on Computers and Accessibility (Assets '04)*. ACM, New York, NY, USA, 86–93. https://doi.org/10.1145/1028630.1028647

Anthony J. Hornof and Anna Cavender. 2005. EyeDraw: Enabling Children with Severe Motor Impairments to Draw with Their Eyes. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '05)*. ACM, New York, NY, USA, 161–170. https://doi.org/10.1145/1054972.1054995

Poika Isokoski, Markus Joos, Oleg Spakov, and Benoît Martin. 2009. Gaze controlled games. *Universal Access in the Information Society* 8, 4 (2009), 323. https://doi.org/10.1007/s10209-009-0146-3

Rob Jacob and Sophie Stellmach. 2016. What You Look at is What You Get: Gaze-based User Interfaces. *interactions* 23, 5 (Aug. 2016), 62–65. https://doi.org/10.1145/2978577

Robert J. K. Jacob. 1990. What You Look at is What You Get: Eye Movement-based Interaction Techniques. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '90)*. ACM, New York, NY, USA, 11–18. https://doi.org/10.1145/97243.97246

Çağla Çığ Karaman and Tevfik Metin Sezgin. 2018. Gaze-based predictive user interfaces: Visualizing user intentions in the presence of uncertainty. *International Journal of Human-Computer Studies* 111 (2018), 78–91.

John King, Julie C Rice, and Andrew Hurley. 2017. Consumer Appeal of Injection IML Packaging vs. Similarly Decorated Glass Jars, Composite Cans, and Metal Cans Using Eye Tracking Technology. *Journal of Applied Packaging Research* 9, 2 (2017), 1.

James F Knight and Chris Baber. 2004. Neck muscle activity and perceived pain and discomfort due to variations of head load and posture. *Aviation, space, and environmental medicine* 75, 2 (2004), 123–131.

Krzysztof Krejtz, Cezary Biele, Dominik Chrzastowski, Agata Kopacz, Anna Niedzielska, Piotr Toczyski, and Andrew Duchowski. 2014. Gaze-controlled Gaming: Immersive and Difficult but Not Cognitively Overloading. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (UbiComp '14 Adjunct)*. ACM, New York, NY, USA, 1123–1129. https://doi.org/10.1145/2638728.2641690

Michael Land. 1992. Predictable eye-head coordination during driving. *Nature* 359 (1992), 24.

Ronald R. Mourant and Charles G. Grimson. 1977. Predictive Head-Movements during Automobile Mirror-Sampling. *Perceptual and Motor Skills* 44, 1 (1977), 283–286. https://doi.org/10.2466/pms.1977.44.1.283

B. Shackel. 1960. Note on Mobile Eye Viewpoint Recording. *J. Opt. Soc. Am.* 50, 8 (Aug 1960), 763–768. https://doi.org/10.1364/JOSA.50.000763

Linda E. Sibert and Robert J. K. Jacob. 2000. Evaluation of Eye Gaze Interaction. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00)*. ACM, New York, NY, USA, 281–288. https://doi.org/10.1145/332040.332445

Bruno Sergi Walter Di Nardo Gaetano Paludetti Fabrizio Ottaviani Stefano Di Girolamo, Pasqualina Picciotti. 2001. Vestibulo-Ocular Reflex Modification after Virtual Environment Exposure. *Acta Oto-Laryngologica* 121, 2 (2001), 211–215. https://doi.org/10.1080/000164801300043541

David J. Ward, Alan F. Blackwell, and David J. C. MacKay. 2000. Dasher&Mdash;a Data Entry Interface Using Continuous Gestures and Language Models. In *Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology (UIST '00)*. ACM, New York, NY, USA, 129–137. https://doi.org/10.1145/354401.354427

David J Ward and David JC MacKay. 2002. Fast hands-free writing by gaze direction. *arXiv preprint cs/0204030* (2002).

Matthias Wille, Lars Adolph, Britta Grauel, Sascha Wischniewski, Sabine Theis, and Thomas Alexander. 2014. Prolonged work with head mounted displays. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers: Adjunct Program*. ACM, 221–224.

AL Yarbus. 1967. Eye movements and vision. 1967. *New York* (1967).