E-Sports is for Everyone: Analysing Eye Movement Patterns in E-Sports Viewers

Elliot McMillan
Clemson University
Clemson, SC, USA
eemcmil@clemson.edu

Evan Rodwell
Clemson University
Clemson, SC, USA
erodwel@clemson.edu

ABSTRACT

E-Sports is a rapidly growing field in the entertainment industry, with over 26.6 million viewers per month from the United States alone. Despite this, the industry does not use interface design techniques to help clarify gameplay for viewers unfamiliar with the content, which results in a large loss of potential viewers. Viewers without prior knowledge might find themselves lost or confused with what is occurring on-screen. This study attempts to identify the eye movement patterns between frequent video game players and infrequent video game players to determine where and how information is being lost when viewing various E-Sports games. By the end of the study, we identify problems in the user interfaces used in two separate E-Sports games, *Hearthstone* and *Overwatch*, that might be improved to increase understanding of the content for infrequent video game players. The study identified the use of empty space, motion, and position on screen of various UI features that could be used and improved on to improve overall game understanding for both groups. This data can be used in further studies on UI to see the effects of the various UI changes to knowledge retention. With enough improvements, it is possible that infrequent video game players can still watch, understand, and enjoy Esports without any prior background knowledge.

CCS CONCEPTS

• **Applied computing → Media arts**; *Consumer products*; *Digital libraries and archives*; • **Human-centered computing →** Empirical studies in accessibility; • **Social and professional topics → User characteristics**.

KEYWORDS

gaze detection, media production, knowledge retention, interface design

ACM Reference Format:
Elliot McMillan and Evan Rodwell. 2021. E-Sports is for Everyone: Analysing Eye Movement Patterns in E-Sports Viewers. In *Proceedings of ACM Conference (Conference'21).* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Conference'21, July 2021, Washington, DC, USA

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

E-Sports is a rapidly growing field in the entertainment industry, with over 26.6 million viewers per month from the United States alone. This is even more impressive when considering that the industry is largest in the Asia-Pacific region. The World Economic Forum reports that over 80 million hours were spent in 2020 viewing E-Sports content on video streaming sites such as Youtube, Twitch, and BiliBili. With such a large industry reporting record profits each year, there is clearly a market for improved user interfaces that allow viewers new to E-Sports to understand and enjoy the content on-screen. However, most of the time, the interfaces used to view E-Sports are identical or nearly identical to their in-game counterparts, which do not take into consideration the need for clarification that viewers new to the game might need. In this study we attempt to identify key regions in the user interface that can be improved upon to increase understanding of the game, and viewing patterns between viewers with various levels of game knowledge to understand how more UI improvements can be made.

Recent studies in eye tracking and digital technologies have begun to examine the eye movement patterns of different groups during video game play, but most of these focus on these eye movement strategies, or differences in gaze pattern. How people watch video games to collect information has been mentioned as a side result in certain studies, but is generally not the main goal of the experimental design. Furthermore, previous experiments rarely studied solely the action of viewing video games and instead focused on playing video games, which is a fundamentally different activity from viewing E-Sports, which do not involve active gameplay. It is possible that eye movement patterns for active play will differ substantially from the movement patterns of those just viewing the same game; thus previous results involving playing games may not be relevant to E-Sports.

The concept for this experiment arose from an observation that most E-Sports games do not significantly change their user interface from the in-game interface, which led to an inability for viewers new to the game to understand the necessary information on screen before it disappeared. The primary contribution of this paper is the identification of problems in the user interfaces of various E-Sports games for viewers who are not familiar with those games. While this is the primary contribution, there are other possible benefits of this experiment, including learning about how different groups watch E-Sports content in general. There could be problem areas that even experienced player of the video games in question have, or this could show that an entire part of the screen is being completely ignored by every viewer. While our study does not test the proposed improvements to those problems, this information can be used in a larger follow-up study to test the effects of changes to those interfaces on knowledge retention.

1.1 Hypothesis

We hypothesize that frequent video game players will have a better understanding as measured by a questionnaire after watching the E-Sports clip as compared to infrequent video game players. In addition, we predict that infrequent game players will have more wild-varying eye movements when information they need is not readily available on-screen, while frequent game players will quickly locate the relevant information while it is available, then have more predictable eye movements afterwards. This will cause longer latencies before fixation for infrequent game players, and longer fixations once the information is found.

Our goal is not to find evidence for the first prediction of the experiment, as this is meant to be an eye-tracking study, but instead focuses on the second prediction. If evidence can be found that infrequent game players have difficulty finding useful information quickly, then changes can be made to the user interface to see if these latencies before fixation can be reduced, thus improving the clarity of the user interface. This information, in addition to other data collected can be used to motivate results beyond our hypothesis. In this case, it would be possible to use the user's gaze data when relevant information is available, if their gaze is *not* directed at the relevant portion of the interface, to change the interface accordingly, and put the relevant information in a more intuitive place for the user.

2 RELATED WORKS

Chen and Tsai investigated hand-eye coordination in children of ages 6-10 and adults while playing virtual reality games [1]. Participants played two games involving reacting to targets that appeared on their screens. One game showed only one target at a time, and participants were to click on the target as soon as possible after it appeared. The other game showed multiple targets at a time, including both targets that the participants were to click on as soon as possible, and others that the participants were to ignore. The authors analyzed participants' gaze patterns as well as their hand motions to identify six strategies used by the participants, differentiated by the order in which they initiated and ended their eye movements and initiated and ended their hand motions. Their results indicate that both children and adults used a combination of hand-eye coordination strategies, but that children had longer latencies and shorter fixation durations than adults, indicating that the children did not yet have adult-level gaze control.

Finke et al. used eye tracking to gauge visual attention in children with and without autism spectrum disorder [2]. Participants were shown a video of *LEGO Marvel Superheroes* gameplay, with the player's face shown in the corner of the video, as is typical for live video game streams. Their results indicate that participants with and without ASD spent similar amounts of time paying attention to the game, and had similar duration of fixation when looking at the player.

Zain et al. used eye tracking to evaluate interface design in educational games [3]. Participants played an educational game which required them to scan text, use checkbox input, and use image navigation and search tools. Their results identified some aspects of the interface that participants found difficult to navigate or understand. They proposed that important aspects of the user interface should be blinking or highlighted to attract user attention, and that instructions should be placed at the center of the screen.

3 EXPERIMENTAL DESIGN

The goal of this experiment is to identify key regions in the user interface of E-Sports games that can be improved upon to increase knowledge retention. While this study could be done as a large-scale study with hundreds of participants, dozens of different games, and additional tests of changes to the user interface, the scope of this study was kept to a smaller scale as a single-semester study, with participants limited to students at the School of Computing at Clemson University. To obtain relevant results, we performed a mixed 2x2 study, using two different groups of users and two different video games, one that has less on-screen information to absorb and is relatively easy to understand, and one that has lots of on-screen content that rapidly changes and may be harder to understand, particularly for participants without prior knowledge of the game. The chosen games based on this criteria were *Hearthstone* as the "easy" game, and *Overwatch* as the "difficult" game. These games both constitute a large percentage of the E-Sports industry, and both have very little changes to the user interface for the E-Sports presentation as compared to the in-game user interface. Figures 1 and 2 show examples of footage from each game, with the areas of interest used in the experiment.

Participants were given a pre-experiment questionnaire to collect background knowledge that might affect the experiment. Then they were presented the stimuli, and were instructed to watch them as they normally would. These were two short, 85-second clips of the selected games. After watching the clips, a short post-experiment questionnaire recorded the participant's opinions on user interface clarity and general knowledge retention.

In addition to questionnaire data, we used Gazepoint technology and PsychoPy software to record eye movements while each participant watched the clips. Number of fixations, fixation location, fixation duration, and eye movements were all recorded. Areas of interest placed on the points of reference in the user interface will record the time of the fixations at various times during the clip, and the overall duration the participant's gaze remained in the areas of interest. Most areas of interest were over key areas of the user interface, such as cards in hand for *Hearthstone* , and both games have an area of interest in the center of the screen to record how long users spend time watching the "focus point" of the game. Notably, both videos included subtitles which participants could view along with the video. The time spent viewing subtitles is relevant to our experiment as it implies the participant prioritized them over presented gameplay.

Using this information, gaze patterns were determined for various times throughout the clips. By considering where fixations occurred and for how long, we can determine key points in the user interface that might need improvement for clarity. If the results of the post-experiment questionnaire showed that the information presented during these fixations was not retained, this would become a relevant result. If the participants fixated for the full time

Figure 1: An example of AOIs over the interface for the E-Sports game *Overwatch*.

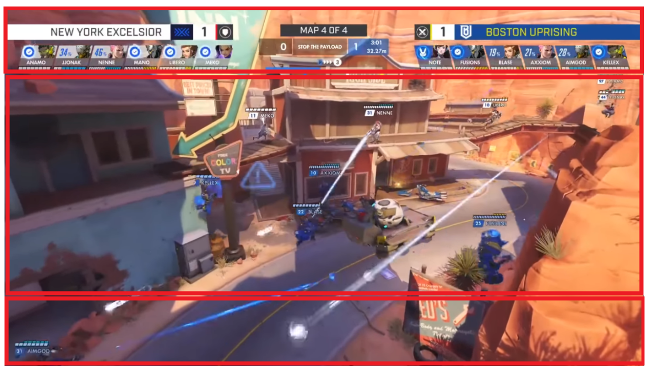

Figure 2: An example of AOIs over the interface for the E-Sports game *Hearthstone*.

while information in the areas of interest were present, this part of the user interface might need to last for a longer time on-screen. If users spent the entire time watching the "focus point" of the screen, perhaps adding more information to unused areas of the screen could improve the user interface's effectiveness at knowledge retention for new viewers, while avoiding obstructing the content for experienced viewers. Results such as these could be used to motivate a larger, more thorough continuation study after this semester.

4 PROCEDURE

4.1 Participants

Participants were recruited from Andrew Duchowski's Eye Tracking Methodology class at Clemson University. 5 participants with varying amounts of prior video game knowledge were recruited. All testing was carried out in Eye Tracking Labs in the Computing Labs of the School of Computing at Clemson University. Participation lasted for no longer than 10 minutes for every participant. Participants ranged in age from 20-25. To ensure participants were familiar with the experiment and to make sure there were no underlying complications the participant may have that could affect the results of the experiment, a pre-experiment questionnaire was given. This questionnaire recorded the participant's age, gender, eye conditions, and if they generally use correctional lenses or contacts to improve their vision.

4.2 Stimuli

Two short clips of prerecorded E-Sports matches were shown to participants, of the games *Hearthstone* and *Overwatch*. Each clip was approximately 85 seconds in duration. Areas of interest were identified beforehand, focused on aspects of the user interface that communicate key facets of the game state. These AOIs are shown in Figures 1 and 2. Even though the experiment intended to collect information on how the user interface and gaze patterns affect game knowledge, the clips were presented with both audio and subtitles, such that the experience was as similar to a regular viewing experience as possible. Information given by the announcers could influence what the viewer understood and where on screen their attention was pulled, and subtitles allowed us to see whether the viewer focused more on clarification of audio over actual gameplay. This is relevant information to collect, as a viewer that focused primarily on subtitles lets us know that the gameplay potentially wasn't as useful to them as the auditory information, and might suggest that more room in the UI should be dedicated to subtitles in Esports.

4.3 Methodology

This was a mixed 2x2 study; participants were sorted into two groups, frequent and non-frequent players, and were shown two different video clips as stimuli. All participants were shown both video clips. Data was also collected on participants' ages and genders.

4.4 Apparatus

Participating students used a 60Hz Gazepoint GP3 eye tracker to complete the required tasks on a 1680 x 1050 monitor. Participants sat in a comfortable position roughly 60 cm in front of the eye tracker, which was then calibrated to ensure distance to the monitor/eye tracker did not affect the experiment. Figure 3 shows an example of this equipment setup. PsychoPy software by Open Science Tools was used to program the presentation of all visual stimuli.

5 RESULTS

5.1 Viewing Patterns

The results we expected to collect for this experiment attempted to confirm our hypothesis, that infrequent game players will have more wild-varying eye movements when information they need is not readily available on-screen, while frequent game players will quickly locate the relevant information while it is available, then have more predictable eye movements afterwards. The results however showed that participant's eye movements followed a predictable pattern between the two groups, with some smaller effects in viewing pattern resulting from prior game knowledge. Both groups ended up viewing the same areas fairly consistently, with those areas generally being the points on screen that were moving during the video. This effect was more pronounced in the *Hearthstone* data, as the *Overwatch* video had a much smaller portion of the screen constantly still.

By analyzing how the motionless parts of the video were viewed, we can understand how the empty/still space was observed between the two groups. The *Hearthstone* video contained a large AOI on the left side of the screen over the face cameras of both players. Since no new information ever showed up on that area of the screen, frequent game players spent at most one single glance at that area, with one participant never looking at the area at all. In comparison, the infrequent game players still viewed this area multiple times. On the top of the screen, another AOI covered the cards in hand of the upper player. This part of the screen moved twice throughout the video, as the player drew and played a card. While the data in that area was relevant to understanding what was occurring in the game, the data was also presented as a moving object in the main "battlefield" AOI. While our hypothesis would suggest that infrequent game players would move their eyes frequently to this area searching for information while frequent game players would look infrequently, locate the information they expected, and look away, the actual results differed greatly. While the frequent game players did infrequently look at this section of the screen then look away after 1 or 2 short fixations, the infrequent game players almost never looked at this section of the screen, with one participant never looking at it at all.

The *Overwatch* video, in comparison, was constantly moving due to the camera following the action of the game in a 3-dimensional environment. Only one part of the interface stayed completely still, the scoreboard in the top AOI of the experiment. While all frequent game player participants viewed this area and had between 2-20 fixations, the infrequent game players never had a long fixation on this section of the screen.

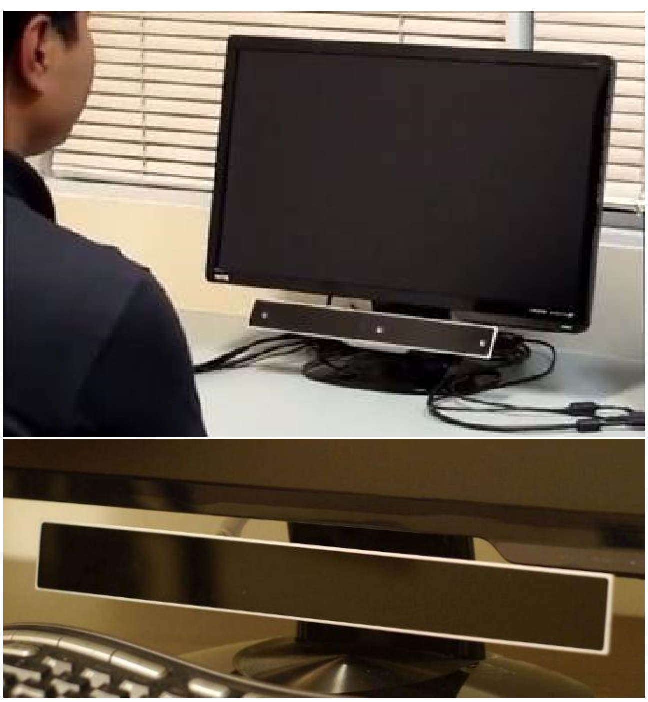

Figure 3: An example of the Gazepoint equipment setup.

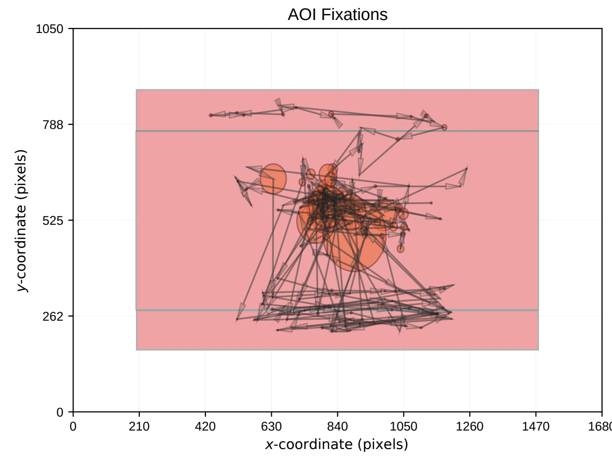

Figure 4: Frequent Game Player 1's Overwatch scanpath.

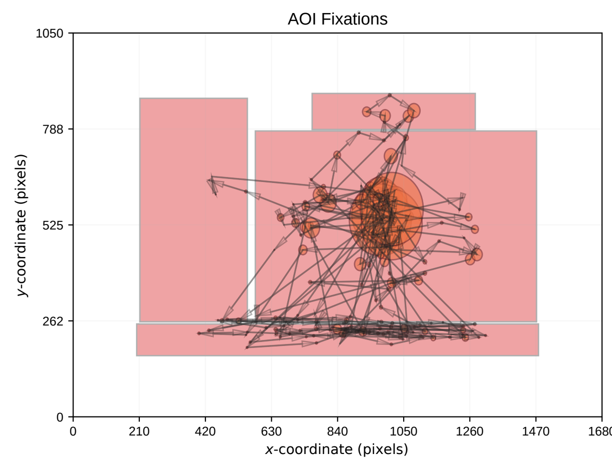

Figure 5: Frequent Game Player 1's Hearthstone scanpath.

The viewing patterns and fixations on the non-still parts of the screen were not observably different between the two groups. Instead, they both followed the primary action on the screen at the time. For the *Hearthstone* video, this was primarily the three points in the middle of the screen that cards appeared in large text during the video. All groups had long fixations on these sections of the videos, with the longest fixations on the very middle card. Aside from these points, participants frequently spent time viewing the subtitles, but rarely fixated on them. The difference in time spent viewing subtitles between the two groups was unnoticeable. For the *Overwatch* video, the camera continually moved its focus towards the center of the action. This resulted in most of the participants spending their time viewing the center of the screen, with large fixations in this area across all participants. The infrequent game players did fixate more frequently to the left and right of this central focus point, but all groups rarely fixated on areas above or below this central point.

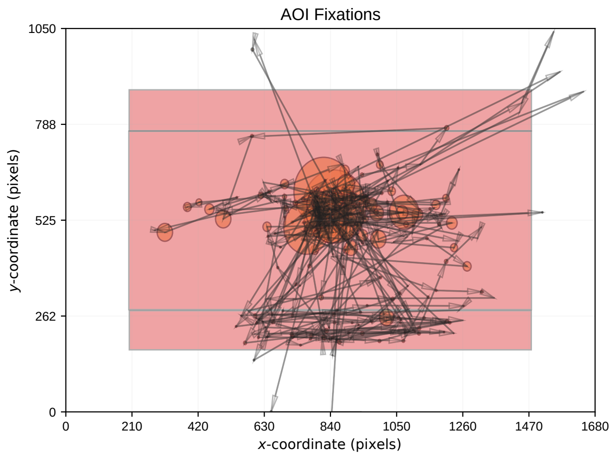

Figure 6: Infrequent Game Player 1's Overwatch scanpath.

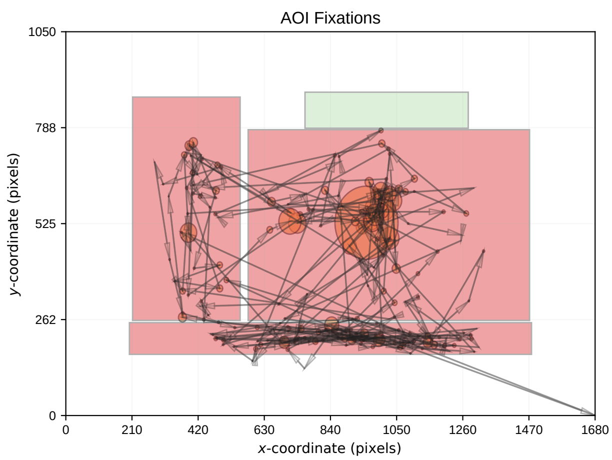

Figure 7: Infrequent Game Player 1's Hearthstone scanpath.

A noticeable effect between all participants and both videos was the amount of almost unobserved or entirely unobserved space in the videos. In the *Hearthstone* video, the rightmost 15 percent of the screen was almost never viewed across every participant. In the *Overwatch* video, the effect was even more pronounced, with the leftmost and rightmost 20 percent of the screen almost never viewed across every participant.

5.2 Survey Results

The survey results showed that the participant's perceived understanding of the information in the video differed from their actual knowledge of the events that occurred. The questions given to the participants ranked from 1 (very difficult/bad) to 7 (very easy/good). Every participant responded that they found it "quite easy" (6) to understand the events in the video, while one non frequent game player responded that it was "easy" (5) to understand. In addition,

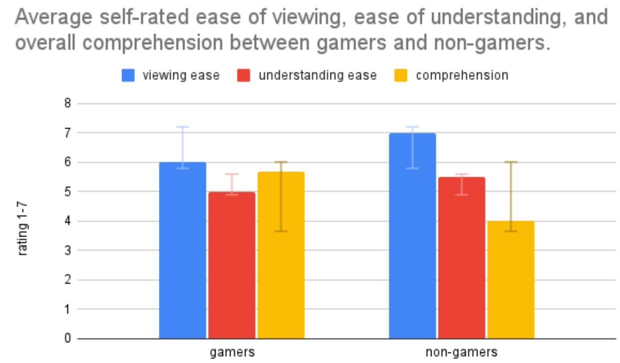

Figure 8: Post-Survey self-evaluation results

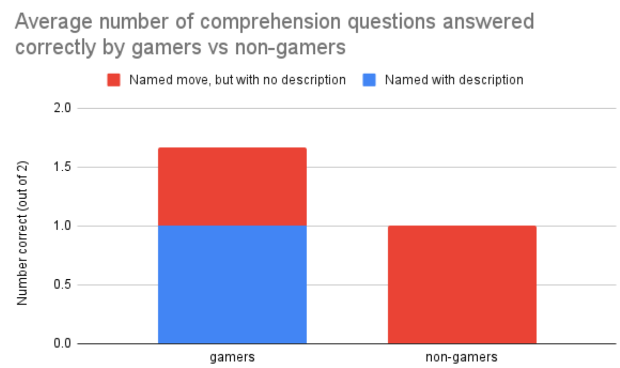

Figure 9: Post-Survey comprehension questions answered

all frequent game players responded that their overall comprehension was "almost good" (5) or above, while infrequent game players responded that their overall comprehension was "neither good nor bad" (4). The questions that asked participants to describe an event from the videos showed that even though the infrequent game players reported only slightly less comprehension of the videos than frequent game players, frequent game players consistently could recall events from both games frequently with descriptions, while infrequent game players could only recall vague descriptions of an event, or nothing at all.

6 DISCUSSION

6.1 Improving Understanding

Infrequent game players perceived their actual comprehension as higher than their comprehension actually was. It is possible this is a bias caused by the viewing ease of E-sports videos, which remained high between both groups. It is also possible that this is a bias caused by the length of our experiment, and if participants were given longer, more comprehensive clips, they would have more reason to improve their understanding and more time to do so. Whatever the cause, our study shows some ways that can be used to improve the clarity of E-sports videos for everyone.

The results of our experiment strongly infer many things about the viewing patterns of people regardless of game experience when viewing Esports gameplay. First, the eye movements of all participants tended to follow the primary movement occurring on screen. This suggests that if there is data that needs to be brought to the attention of a viewer to improve their understanding, the primary design change to implement this should be to involve motion to the way that data is presented, rather than focusing on the size or location of the data. While changing the size and location of the data can also improve if the data is viewed, the more effective method is to add motion.

Another method of ensuring that the viewer sees relevant information is to put that data in the center of the screen. While putting the information closer to the center may improve the chance of the viewer fixating on it, the most likely locations for the viewer to see the information is either directly in the center or just to the left or right of this area. Our data suggests that putting the data above or below the center of the screen is more effective than putting it in empty white space near the screen's edge, but is not as effective as the locations horizontally close to the center.

6.2 Improving UI

There are many changes that could be made to the visual presentation of both games and their user interfaces to increase player knowledge. The most obvious change involves the changes to empty spaces on the screen. Where data showed up for short periods of time on-screen, it may be helpful to keep that information visible on the parts of the screen people rarely ever view or use.

For *Hearthstone*, this would include the rightmost 15 percent of the screen, in addition to the corners of the screen. Since our data also suggests that relevant data should be presented in the center of the screen, changing cards played to appear in the center of the screen, then use motion to move to one of the infrequently used empty spaces identified by our experiment could drastically improve understanding for infrequent game players while not obstructing game view for frequent game players.

For *Overwatch*, this would include the leftmost and rightmost 20 percent of the screen, or the scoreboard at the top of the game interface. Since there was motion on the left and rightmost parts of the screen throughout the entire video that contained gameplay, and the camera focused on that main gameplay very well, the leftmost and rightmost parts of the screen still contained relevant information that was at least viewed in the periphery of a viewer's vision. The scoreboard however was motionless, and barely used by either group despite containing useful game information. Improvements could be made to show current events more clearly on this part of the screen such as abilities and deaths, with some sort of animation to add motion and draw the viewer's attention. Since infrequent game players never had a long fixation on this section of the screen, this suggests that the data given there was not useful to them at all without more context. Adding additional clarity to this scoreboard through size, motion, and vocally bringing attention to these statistics could make this section of the screen more useful to all players, which is important as it already takes up a large percentage of the screen.

6.3 Subtitles

Even though the video was presented alongside audio, a large portion of all participant's time across both groups was spent reading the subtitles. It is possible that this could be the result of two aspects of the experiment:

Firstly, participants may have needed clarification on what the audio was stating. This is the usual reason that subtitles are used, but an experiment with the presence of subtitles as an independent variable could better gauge how much the overall understanding of what was occurring on-screen was affected by their presence. It is possible that participants could even improve their understanding without subtitles, as they would spend more time viewing the gameplay.

Secondly, participants may have focused on the moving parts of the screen, which was already a noticed pattern in the main viewing patterns of both groups. This is even more likely due to the fact that more time was spent viewing the subtitles in the *Hearthstone* video over the *Overwatch* video, even though the subtitles are larger in the latter. Hearthstone had a larger amount of time with no movement on the screen other than the subtitles, which may have caught the participant's attention.

7 CONCLUSION

We conclude that our initial hypothesis was only partially supported. Our theory that frequent game players would quickly locate any relevant information while it is available, then have more predictable eye movements afterwards was partially supported. The frequent game players did in fact view the informational portions of the screen more frequently than the infrequent game players, and also avoided the sections of the screen that lacked useful information more than the infrequent game players. The rest of the hypothesis, that infrequent game players would have more wild-varying eye movements when information they need is not readily available on-screen, and that this would cause longer latencies before fixations for infrequent game players and longer fixations once the information is found, was not supported. Instead, both groups viewed the experiment prioritizing the center of the screen, and motion occurring around that point. Instead of having more varying eye-movements, the infrequent game players still viewed similar parts of the screen to frequent game players due to this presence of motion on relevant information. This was more profoundly noticeable in *Hearthstone*, as the game contained much less constant motion and a large amount of still space.

Our experiment also shows many points on screen that neither group frequently viewed, which can be improved upon to increase clarity for infrequent game players without getting in the way of the viewing patterns of frequent game players. Through use of knowledge of frequently viewed space, infrequently used space, and the viewing patterns of frequent and infrequent game players provided by this study, changes to the user interface to improve clarity can be made more effectively for both groups.

Further studies could be performed on the impact of our suggested changes to the user interface, as well as the effects of subtitles on viewing patterns of E-sports gameplay.

REFERENCES

[1] Yuping Chen and Meng-Jung Tsai. 2015. Eye-hand coordination strategies during active video game playing: An eye-tracking study. *Computers in Human Behavior* 51 (2015), 8–14. https://doi.org/10.1016/j.chb.2015.04.045

[2] Erinn H. Finke, Krista M. Wilkinson, and Benjamin D. Hickerson. 2017. Social Referencing Gaze Behavior During a Videogame Task: Eye Tracking Evidence from Children With and Without ASD. *Journal of Autism and Developmental Disorders* 47, 2 (01 Feb 2017), 415–423. https://doi.org/10.1007/s10803-016-2968-1

[3] Nurul Hidayah Mat Zain, Fariza Hanis Abdul Razak, Azizah Jaafar, and Mohd Firdaus Zulkipli. 2011. Eye Tracking in Educational Games Environment: Evaluating User Interface Design through Eye Tracking Patterns. In *Visual Informatics: Sustaining Research and Innovations*, Halimah Badioze Zaman, Peter Robinson, Maria Petrou, Patrick Olivier, Timothy K. Shih, Sergio Velastin, and Ingela Nyström (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 64–73.