MEASURING THE EFFECT OF TASK-IRRELEVANT VISUALS IN AUGMENTED REALITY

by

Allison Hopkins

A Thesis

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Master of Science

Computer Graphics Technology

West Lafayette, Indiana

May 2019

THE PURDUE UNIVERSITY GRADUATE SCHOOL
STATEMENT OF COMMITTEE APPROVAL

Dr. David Whittinghill Chair

School of Computer Graphics

Technology Dr. Austin Toombs

School of Computer Graphics

Technology Dr. Eyal Ofek

Microsoft Corporation

Approved by:

Dr. Colin Gray & Dr. Nicoletta Adamo

Head of the Graduate Program

Dedicated to my family

ACKNOWLEDGMENTS

I would like to thank my sister, Dr. Lauren Hopkins, my parents, and my friends for encouraging me to keep going and let me bounce ideas off them. I would also like to acknowledge my advisor, Dr. David Whittinghill for continuing to challenge me, and Dr. Robert Proctor who showed me the importance of human factors in designing future interfaces. Finally, I would not have this opportunity if it were not for my manager, Dan Rhine, who made it possible for me to get this degree in the first place.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY

Augmented reality (AR) – "An enhanced version of reality created by the use of technology to overlay digital information on an image of something being viewed through a device (such as a smartphone camera)" (Merriam-Webster Incorporated, 2018).

Cost of concurrence – "The difference between the performance level on a task when it is performed alone versus when it is performed with another task to which no attentional resources are devoted" (Proctor & Van Zhandt, 2008, p. 560).

Divided attention – "The act of focusing attention on several sources of input at once" (Proctor & Van Zhandt, 2008, p. 561).

Head-mounted display/Helmet-mounted display (HMD) – "A display mounted on a helmet worn by a person that can be seen no matter where the person is looking" (Proctor & Van Zhandt, 2008, p. 563). A head-mounted display is similar to the helmet-mounted display except that it sits on the user's head like a visor instead of completely encasing the head. For the purpose of this research, "HMD" will refer to head-mounted display unless otherwise specified.

Heads-up display (HUD) – "A display on the windshield of an aircraft, automobile, or other vehicle that allows the operator to read the display without having to direct his or her gaze away from the outside world" (Proctor & Van Zhandt, 2008, p. 563).

Mobile augmented reality (MAR) – "The employment of AR concept in mobile devices, which means increasing user perception about the real world with virtual elements over and synchronized to that real world, visualized through mobile devices" (Santos et al., 2016).

Performance operating characteristic (POC) – "A performance operating characteristic is an examination of how performance measures on two (or more) tasks are interrelated as the allocation of resources among these tasks is changed" (Norman & Bobrow, 1976).

Situational awareness – "Consciousness of the objects in the environment, what they mean, and their future status" (Proctor & Van Zhandt, 2008, p.568).

Virtual reality (VR) – "Virtual reality attempts to replace a user's perception of the surrounding world with a computer-generated artificial three-dimensional (3D) VE (virtual environment)" (Hou et al., 2013).

ABSTRACT

Author: Hopkins, Allison, C.
MS Institution: Purdue
University Degree Received:
May 2019
Title: Measuring the Effect of Task-Irrelevant Visuals in Augmented
Reality Committee Chair: Dr. David Whittinghill

Augmented reality (AR) allows people to view digital information overlaid on to real-world objects. While the technology is still new, it is currently being used in places such as the military and industrial assembly operations in the form of ocular devices worn on the head over the eyes. Head-mounted displays (HMDs) let people always see AR information in their field of view no matter where their head is positioned. Studies have shown that HMDs displaying information directly related to the immediate task can decreased cognitive workload and increase the speed and accuracy of task performance. However, task-irrelevant information has shown to decrease performance and accuracy of the primary task and also hinder the efficiency of processing the irrelevant information. This has been investigated in industry settings but less so in an everyday consumer context. This study proposes comparing two types of visual information (text and shapes) in AR displayed on an HMD to answer the following questions: 1) when content is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least? And 2) When presence is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?

CHAPTER 1. INTRODUCTION

A technological race is currently occurring as companies aim to be the first to produce a commercially available ocular wearable for everyday use. A press release from the Consumer Technology Association (CTA), who showcases new technology at its annual Consumer Electronics Show (CES), projected that augmented reality (AR)/virtual reality (VR) "will sell 4.9 million units in the U.S. in 2018 (a more that 25 percent increase from 2017) and see U.S. revenues of $1.2 billion (almost 10 percent year-over-year growth)" (Cassagnol, 2018, para. 2) with AR ocular devices growing within the next five years.

Such devices aim to have an integrated heads-up display (HUD) which allows the user to see the real world overlaid with digital information. The successor of VR, AR interfaces are moving out of controlled environments (e.g., factory operations, military training rooms, hospital operating rooms, etc.) and beginning to show up in everyday life (e.g., mobile phone games, vehicle HUDs, Google Glass, etc.) as well as both indoors and outdoors (Martínez, Skournetou, Hyppölä, Laukkanen, & Heikkilä, 2014). However, one of the biggest concerns associated with AR is figuring out how to ensure the safety of consumers by reducing user distraction when using AR HUDs in everyday situations. This concern is paramount as human-computer interaction within altered reality environments continues to grow.

1.1 Statement of the Problem

It has been established in several studies (Bark, Tran, Fujimura, & Ng-Thow-Hing, 2014; Haiduk, 2017; He, Choi, McCarley, Chaparro, & Wang, 2015; Hou et al., 2013; Mustonen, Berg, Kaistinen, Kawai, & Häkkinen, 2013; Rane, Kim, Marcano, & Gabbard, 2016; Woodham, Billinghurst, & Helton, 2016; Yeh, Merlo, Wickens, & Brandenburg, 2003) that AR HUDs can both help and hinder the performance of a task but it all depends on how that information is displayed along with the purpose of it. Studies of AR devices such as Google Glass (He et al., 2015) and Shimadzu Data Glass (Mustonen et al., 2013)) concluded that the primary task (i.e. driving or walking) in conjunction with an AR overlay did indeed distract the participants by causing loss of situational awareness and driving/walking deviations.

However, it is important to note that the information in these cases were not task related, meaning it does not help the user by providing information associated with their primary task (e.g., walking down the street to a specified destination). Such an overlay of information can take up precious cognitive resources and obscure real-world objects. Even though task-irrelevant information has shown to distract under most conditions, difference modalities such as audio have been explored to mitigate the distractor factor. However, audio messages and alerts are only optimal if immediate action is required and the notification will not need to be recalled later (Proctor & Van Zhandt, 2008, p. 194). Alternately, visual messages are the opposite and are best used when no immediate action is required and notifications will need to be remembered later (Proctor & Van Zhandt, 2008, p. 194).

The best way to display task-irrelevant information in AR is not to display it all. However, this is not feasible, as people will continually seek out information with or without AR. For example, with modern smartphones, 74% of drivers between the ages of 17 and 29 admitted to texting while driving on a regular basis and less than 10% accessed the web while driving (Cook & Jones, 2011). With ocular wearables and other smart HMDs forecasted to hit the consumer market in the very near future (Phelan, 2019), identifying and designing around these distractor factors will be paramount for the safety of the users.

1.2 Research Questions

As human cognition is limited, attention cannot be split evenly and a cost must be paid to either the primary task or the distracting non-task visual resulting in a decrease of attention of one or the other. It is known that adding a secondary, irrelevant task will slow down completion of the primary task, but is there a difference in how much it affects the primary task based on how it is displayed? This study seeks to answer the following research questions:

1. Does the format of notification (text or shapes) affect the performance of a non- related primary task and, if so, which affects performance the greatest?
2. When content is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?
3. When presence is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?

The hypothesis (H_1) states that one of the notifications (text or shape) will have a significantly greater effect than the other on the performance of the primary task while the null hypothesis (H_0) states that neither of the notifications will have an effect on the efficiency on the primary task.

1.3 Significance

With consumer technology trending towards augmented wearables (Martínez et al., 2014), how users interact and perceive information within these devices will affect the future of human- interface design. Past research is at an impasse; on one hand, AR devices can be less distracting than other technology (He et al., 2015; Sawyer et al., 2014) but it is still considered distracting nonetheless. On the other hand, studies such as those done by Bark and colleagues (2014) concluded that interfaces can be designed so that users will suffer no distraction of their primary task. In addition, AR interfaces such as those used for aviation can actually enhance situational awareness by reducing "heads down" time when looking at critical instruments inside the cockpit (Haiduk, 2017).

Sabelman and Lam (2015, p.50) pointed out that this disparity could be related to the way the information is presented: aviation heads-up displays (HUDs) "typically shows information in a highly symbolized and minimalistic way, with little text and no images of people (…)." Thus, the possibility of designing an interface to lessen distraction is worth exploring.

The studies are few in number that have used head-mounted displays (HMD) in investigating distraction as opposed to mobile augmented reality (MAR) applications which could be due to the recent development of commercially available AR HMDs outside of industrial or military environments (Martínez et al., 2014). It is important to differentiate between HMDs and MARs because the former is hands free while the latter is not. Performing a physical task with an AR overlay is more easily done with the device attached to one's head rather than holding it up by hand.

Something that has not been explored, however, is the cost of concurrence between these two tasks: that of a primary motor task using visual resources and an independent non-task related visual shown in AR. In addition, comparing text versus minimalistic graphics in the form of shapes can show which distracts the user more. By analyzing this data, the type of information presented to the user can be a strong indicator of when the primary task begins to suffer at the cost of the non-task visual. This information can then be used to inform and optimize future AR HMD interface designs so users can interact with real objects while having minimal distractions from non-task related AR visuals.

1.4 Statement of Purpose

This study will address the issue of the level of distraction of the type of notifications by using a timed constant task and hit/false alarm *d'* assessment used in signal detection theory (Mustonen et al., 2013). With the advent of smart glasses and other ocular wearables, this research will focus on HMDs and not MAR applications.

The purpose of this research is to determine the attentional cost on a primary physical task when non-task related information is presented at the same time. Since information can be present in different forms, the proposed research will test the level of distraction only on visual notifications. With this, the level of distraction will be analyzed and guidelines for presenting information in an AR HUD can be created and offered as "best practice" for future interface designs.

1.5 Assumptions

Because this experiment will be conducted within-subjects, the following assumptions have been made:

1. The HoloLens itself (e.g., weight, headband tightness or other hardware factors) is not a detriment to participants' performance.
2. Participants are honest about their vision and have brought appropriate corrective lenses to perform the experiment.
3. Participants are familiar with the shapes/text shown on the HoloLens and can read English text.
4. The results of this experiment conducted in a controlled environment can be applied to commercialized consumer ocular wearables in a naturalistic environment.

1.6 Limitations

The following are known limitations to the proposed study:

1. Possible lack of participants can negatively impact the significance of the statistical analysis.
2. Use of available participants by way of convenience sample may not apply to a larger population.
3. Participants may not be honest in reporting visual impairments.

1.7 Delimitations

The following are known delimitations to the proposed study:

1. Participants in the study will be men and women the West Lafayette, Indiana and Lafayette, Indiana area. Therefore, the sample might not be applicable to a much wider population including education level, socioeconomic status, or familiarity with certain technological devices.
2. This study will only address those with normal 20/20 vision with or without corrective lenses. The effects of the experiment on those with color blindness or other vision impairment is outside the scope of this project.

This experiment will only be utilizing the Microsoft HoloLens due to equipment availability and the fact that it is the only untethered AR HMD accessible to the researchers at Purdue University.

1.8 Summary

With the advent of ocular wearables, a proactive approach to designing AR HUD interfaces can identify potential safety concerns that are present in current MAR applications and other HMDs. However, the goal of MARs and other portable devices is to grab and hold on to the attention of the user that disrupts any primary task that person might be engaged in (e.g., walking, driving, reading etc.). It is therefore important to understand how overlaying information onto the real world affects a person's ability to perform such tasks.

The proposed experiment will address the following research questions:

1. Does the format of notification (text or shapes) affect the performance of a non- related primary task and, if so, which affects performance the greatest?
2. When content is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?
3. When presence is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?

The hypothesis (H_1) states that one of the notifications (text or shape) will have a significantly greater effect than the other on the performance of the primary task while the null hypothesis (H_0) states that neither of the notifications will have an effect on the efficiency on the primary task.

CHAPTER 2. REVIEW OF LITERATURE

2.1 Current HUD Uses in Industry

The first mainstream use of a HUD began in the military and was used by fighter pilots to view critical data without having to continuously look in and out of the cockpit (Vogel, Schultz, & Schultz, 2004). The integration of HUD technology into a head-mounted display (or helmet- mounted display) (HMD) allowed fighter pilots to not only see this data but to also have the data move with their heads, always keeping the information always in their field of view. This kind of display is beneficial for pilots because it reduces "head-down" time, or the time pilots spend with their heads down looking at instruments, maps, and other information inside the cockpit (Haiduk, 2017).

This display technology is also beneficial outside of the military. Haiduk (2017) researched the effects of smart glasses (glasses with overlying AR information) in a general aviation setting and found that pilots made less tracking errors with the glasses than with normal head-down methods of looking at instruments and maps. However, when analyzing spatial awareness between the two groups, there was no significant difference between reaction times although the number of signals missed was less in the smart glasses group (Haiduk, 2017). The reason for this was attributed to the increased time spent scanning outside the cockpit (as pilots are trained to do) but the research design could also have been a factor; the pilots were told to press the "push-to-talk" button when the stimulus was noticed which is something pilots may not normally do as the "push- to-talk" button is reserved for communication only (Haiduk, 2017).

Outside of aviation, another area that requires consulting information from a source other than on the object itself is in assembly operation. Operators often have to consult traditional paper manuals for assembly instructions that can range from simple to complex and may at times be unclear or redundant (Hou et al., 2013). Not only does this method take a substantial amount of time to complete, but the process is subject to human errors when the operator repeatedly switches from reading the instructions to physically putting together the object (Hou et al., 2013). Research using AR to train operators in assembly tasks, however, showed that not only was learning increased (as opposed to relying on memory or the paper manual) but assembly times, errors, and perceived cognitive load decreased (Hou et al., 2013). Having useful information overlaid on top of the physical environment is similar to resolving the pilot's "heads-down" issue; the operator does not have to constantly consult a physical manual and take eyes off of the artifact being assembled in order to piece it together correctly.

Though AR has been extremely valuable for improving situational awareness for pilots and industrial assembly operations, the opposite is true for certain everyday situations. Since consumer AR is still in its infancy, one of the most popular applications of use is for entertainment. It was reported that several users of Pokémon Go, one of the most popular and widely available AR games for mobile devices, played the game while driving or walking and lost sense of their surroundings, resulting in accidents either to themselves or to others (Ayers et al., 2016; Joseph & Armstrong, 2016). One particular case in where a man playing the game while driving resulted in a rollover while another incident landed a 58-year-old woman in the hospital after swerving to avoid a pedestrian who had walked into the road while catching Pokémon (Joseph & Armstrong, 2016).

An analysis of "tweets" via Twitter in 2016 indicated a similar trend: 33% revealed that a driver, passenger, or pedestrian was distracted by the game, 18% revealed self-reported use of playing the game while driving, and 11% of those playing were passengers in motor vehicles (Ayers et al., 2016).

Though this can be attributed to the same "heads down" effect that pilots suffer from (i.e., people looking down at their smartphones), the larger issue potentially stems from a competition of cognitive resources and loss of situational awareness.

2.2 The Importance of Task Relevance

When more than one task is being attended to at the same time, it becomes a divided-attention task and competes for limited cognitive resources (Proctor & Van Zhandt, 2008, p. 242). The dual-task paradigm is well suited to determine the point of cognitive overload and can be used to predict human performance (Proctor & Van Zhandt, 2008, p. 234).

Several experiments have been conducted that addresses the distraction problem of motor control and cognition. Mustonen, Berg, Kaistinen, Kawai, & Häkkinen (2013) used a monocular HMD (where the augmented interface was only situated in front of one eye) and had participants perform the dual-task of walking and recall of information displayed on the interface. Walking performance did indeed decrease, which was measured by speed and path overruns while keeping a timed pace (Mustonen et al., 2013). But perhaps the most significant result was that when the participants tried to control their steps while recalling the information on the interface, the performance of both tasks suffered: the walking pace worsened along with visual vigilance and memory recall (Mustonen et al., 2013).

Another example of the effects on AR on a physical task is a study done by Woodham, Billinghurst, and Helton (2016) where they measured rock climbing performance and word recall using an AR HMD. Their results also showed a decrease in motor and cognitive functions in the form of impaired/slowed climbing and decreased word recall (Woodham et al., 2016). This impairment was attributed in part to the “visual clutter” the AR interface caused the participants whose primary task was climbing (Woodham et al., 2016).

Overlying information on physical objects does not just obstruct the real world, it also obscures information people gather from physical objects and their surroundings. For everyday situations, central and peripheral vision become are extremely important when processing information. Mustonen et al. (2013) pointed out that tasks that utilize central vision also affect how information is processed in the periphery in a dual-task paradigm. By placing a HMD over the participants’ eyes (or over one eye, in the case of the Google Glass), the device is limiting or obstructing the peripheral vision of the wearer which can change how information is processed (Mustonen et al., 2013). To date, there is no commercially available AR HMD that includes peripheral vision in its design; all current devices have the interface situated in the wearer’s central vision only.

One explanation of the deterioration of both tasks is Wickens’s multiple-resource model, which states that two visual tasks conflict with one another because they draw from the same resource pool (Wickens, 2008). This points to Wickens’s visual channels that are defined as focal (the ability to read and recognize objects) and ambient (body awareness involved in orientation of movement and coordination) (Wickens, 2008). Therefore, when someone is engaged in a visual task such as reading a notification on an AR HMD and walking at the same time, one task will suffer at the cost of the other. This supports the theory of resource-limited cognitive ability where, in a dual-task paradigm, the more cognitive resource a person has to focus on a task, the more performance will improve on that particular task (Norman & Bobrow, 1976).

However, when another task is introduced and both are being attended to simultaneously, the relationship becomes more complicated. If two tasks meet the criteria of being "resource-limited," a finite amount of resources a unit has to perform a function, the tradeoff in performance can be measured with a performance operating characteristic (POC) curve which will show an increase in performance of one task and a decrease in the other (Norman & Bobrow, 1976). Pashler (1994) equated this to a bottleneck of resources called the psychological refractory period (PRP) effect when two tasks compete for a person's attention resulting in a brief processing delay or complete impairment. This assumes, however, a single channel for processing information for two tasks requiring the same resources (cognitive, physical, auditory etc.). When coupled with Wickens's multiple resources theory (2008), this makes logical sense since someone performing an auditory task coupled with a visual task will do better than someone performing a visual spatial and memory recall task.

Deciding which modalities to use together has such an extreme impact on designing human-computer interfaces; displaying information or asking a human operator to perform a task (or several tasks) while simultaneously giving that person instructions can either help improve performance or overload the operator's ability to comprehend and complete the task.

2.3 Displaying Information

With the benefits of AR HUDs well documented in the aviation industry, the automotive industry sought to see if the same benefits could be applied to driving vehicles on the road. The effects on AR HMDs and user distraction were investigated using the Google Glass (an AR HMD which displays information just in front of the user's eye), the results of which showed that while the Google Glass was less distracting than texting on a smartphone (resulting in a decrease in lane, steering wheel, and heading deviations), it was still distracting nonetheless (He et al., 2015; Sawyer et al., 2014). Sawyer and colleagues (2014) suggested that cognitive interference played a key role in some of these findings, making the point that when someone is presented with information, the technology delivering the information makes less difference than the fact that new information is placing cognitive demand on working memory. Though this is specifically aimed at text-based messages, factors other than the format of the information are also important including the content of the message itself, the processing of the brain decoding the message, and the subsequent effect of the impact of the messages on the reader

(Sawyer et al., 2014). However, a critical analysis of this suggestion yields that reading a stop sign, highway sign, or other text-based road marker also meets these criteria but does not seem to impair the drivers' cognitive functions. For example, a closed expressway exit might be unexpected but can easily be accommodated for by following detour signs.

Driving a vehicle can be considered a stressful and potentially cognitively demanding task due to many unexpected variables (i.e., other vehicles on the road, pedestrians, unexpected road hazards such as potholes etc.) but, then again, so can aviation. The difference between these two scenarios, however is twofold: 1) the purpose of the information presented and, 2) the way it is presented along with the training of those encountering these scenarios.

Firstly, in the prior road sign example, any sign the driver encounters (excluding billboard advertisements in this case) is informationally relevant to the task of driving. The same holds true for AR HMDs used in aviation; the information given to the pilot is only relevant to the immediate task. While "task relevant information" depends on the situation, it means information that directly relates to the task at hand for assisting the person performing the task by providing information. Several studies (He et al., 2015; Sawyer et al., 2014) that used AR HMDs in the context of driving have presented information that are not task relevant and the conclusions have been that while the devices can be less distracting than others, they still distract. When someone puts on an HMD and AR information is placed over the real world, it creates visual clutter in an environment where there previously was none. Pilots can use HMDs effectively because they are trained to do so but in other scenarios this becomes more problematic.

In a study done by Yeh, Merlo, Wickens, and Brandenburg (2003), the effects of displaying task-irrelevant data via a HMD were investigated and found that this overlay of unnecessary information did indeed hinder the user's ability to detect targets even though it reduced the need to scan for the necessary information. Not only that, but when task-irrelevant data was increased, the user's ability to perform the task decreased (Yeh et al., 2003). This "cost of clutter" outweighed the potential benefits of scanning provided by AR HMDs (Yeh et al., 2003). Therefore, any visual elements that are not relevant to the task at hand is detrimental to cognitive processes as seen with the climbing study (Woodham et al., 2016).

The experiment done by He (2015) is another example as participants had to respond to general text messages on either a smartphone or Google Glass using Google Voice. Since the text messages had no relevance on the primary driving task, driving performance suffered for both devices but less so for the Google Glass (He et al., 2015). Other vehicular AR HUDs have been explored but further research needs to be done to determine the effectiveness of the display.

One such example is a study done by Bark, Tran, Fujimura, and Ng-Thow-Hing (2014) who pointed out that, designed correctly, AR HUDs can assist people with maintaining situational awareness in the context of vehicular navigation. They showed participants video clips of different driving scenarios with overlaid AR graphics of an arrow and short text added in postproduction (e.g., “turn right”) and measured the time it took for participants to recognize where to turn or take a particular exit (Bark et al., 2014). Recognition time of the turns was faster when using the AR visuals in five out of eight scenarios with the non-AR video being quicker in identification due to obstruction of physical objects like overpasses (Bark et al., 2014). However, participants were not manually performing the driving task themselves so it cannot be accurately concluded that the AR display eased their cognitive load instead of adding to it.

In a similar context, a design study by Rane, Kim, Marcano, and Gabbard (2016) analyzed the difficulties of novice non-U.S. native drivers and built an interface to assist them with maintaining awareness of their surroundings while easing cognitive load. The researchers were careful to consider “information density”, or the amount of information shown to the participants so as not to overload them (Rane et al., 2016). A usability test was conducted by two heuristic experts using video footage and a simulator, the results of which were found to be promising in improving decision making (Rane et al., 2016). However, this was a proof of concept only and no actual non-native participants tested the interface to confirm these results (Rane et al., 2016).

Secondly, how the information is presented is just as important as to what is presented. The participants in the Google Glass driving and climbing studies (He et al., 2015; Sabelman & Lam, 2015; Woodham et al., 2016) were shown text based messages and tasked with dividing their attention between reading what was shown and performing a physical task. However, as was previously discussed, both pilots (Haiduk, 2017) and drivers in certain simulated conditions (Bark et al., 2014; Rane et al., 2016) can benefit from overlaid information. Sabelman and Lam (2015, p. 50) pointed out that this disparity could be related to the way the information is presented: aviation HUDs “typically shows information in a highly symbolized and minimalistic way, with little text and no images of people (…).” This was something Rane and colleagues (2016) had considered when designing their experimental AR interface; the shape, size, position, and colors of notifications along with accompanying text, resulting in a design that displayed only relevant information as clearly as possible.

Task irrelevance and information display type are two variables that are often paired together but the relationship between them has not been explored in great detail. While it is accepted that in a dual-task paradigm one cannot divide attention between tasks without a cost to the other (Norman & Bobrow, 1976), the effect of the type of information of non-task visual elements has not been measured.

What Rane and colleagues (2016) did by combining visual icons and text instructions was interesting in regard to interface design for an experiment investigating the effectiveness of navigational aids. This “dual coding” of information (presenting it in more than one form) seemed not to hinder the performance of the participants. In addition to this, the prototype not only had information displayed symbolically and textually but auditory as well in the form of “turn right/left”, “stop”, or “slow down” (Rane et al., 2016). This supports the work done by Mayer and Moreno (1998) who analyzed the effects of learning with a visual animation paired with text versus the animation paired with auditory narration. While learning is different from simple signal tests, the modalities for relaying information can be comparable. It was concluded that retention was higher with the participants who viewed the animation in conjunction with the auditory narration due to the dual-processing theory which states that learning will be greatest when the information uses separate modalities (Mayer & Moreno, 1998). Rummer, Schweppe, Fürstenberg, Scheiter, and Zindler (2011) made the point that because text and pictures take up the same visual modality, it is more difficult to focus on both at the same time.

One study done by Whittinghill and Herring (2017) explored animation and text versus just animation instructions in game design and found that the animation only group performed the same and in some cases better than the animation and text group. This also points to the dual coding theory that additional visual instructions can conflict with each other and, by removing one of the conflicting variables, performance can be increased.

However, dual coding of information for interface design has not been extensively researched and is especially lacking in the area of AR. In addition, it cannot be assumed that research done in the area of games or two-dimensional (2D)/three-dimensional (3D) computer graphics will be applicable for AR interface design. The proposed study will therefore investigate different types of visual AR information and the effect these types have when performed in conjunction with a primary cognitive task.

With the exception of Haiduk's (2017) smart glasses, it should be noted that the devices used in the aforementioned experiments have either been monocular HMDs (like the Google Glass), handheld mobile augmented reality (MAR) applications (like Pokémon Go), simulators with post-production graphics, or simply AR screens in place of HMD technology. With the recent release of the Microsoft HoloLens in 2016 (Kipman, 2016), researchers now have a self-contained untethered AR HMD which is the closest thing to ocular wearables that is currently available. What the previous studies lack is the use of such a device to accurately gauge cognitive workload of a dual-task paradigm for an AR interface.

CHAPTER 3. METHODOLOGY

3.1 Participants

Participants were recruited from the Purdue University campus via flyers and emails and answered a questionnaire regarding current smartphone usage and any vision impairments. Participants were instructed to bring corrective eyewear if their natural vision was not 20/20. A $25 Amazon gift card was given to the participant who identified the most signals correctly and scored the highest in the matching game.

7 females and 17 males (25 participants total) took part in the experiment. However, due to technical glitches where one subject's responses were not recorded, only the data from the last 24 participants were analyzed. The score of this subject was used instead as pilot data for the statistical measures of the correct and incorrect game scores. Of these 24 participants, 54.17% were aged 18- 24, 20.83% were aged 25-34, 16.67% were aged 35-44, and 8.33% were aged 45 and over (see Figure 1). The average daily smartphone use was reported to be between 3-4 (37.5%) and 5-6 hours (37.5%) with 16.67% being between 1-2 hours and 8.33% being between 6-10 hours (see Figure 2). The purpose of this smartphone usage was mixed: 13.13% reported using it for sending/receiving text messages, 13.13% for sending/receiving emails, 7.5% for playing games, 13.75% for accessing the internet, 12.5% for listening to music/watching videos/streaming content, 11.88% as a navigational aid or other location-based service, 13.13% for social media, 11.88% for reading news/e-books/or other digital content, and 3.13% of miscellaneous use including taking photos/videos, tracking sleep/steps, and as an alarm clock (see Figure 3).

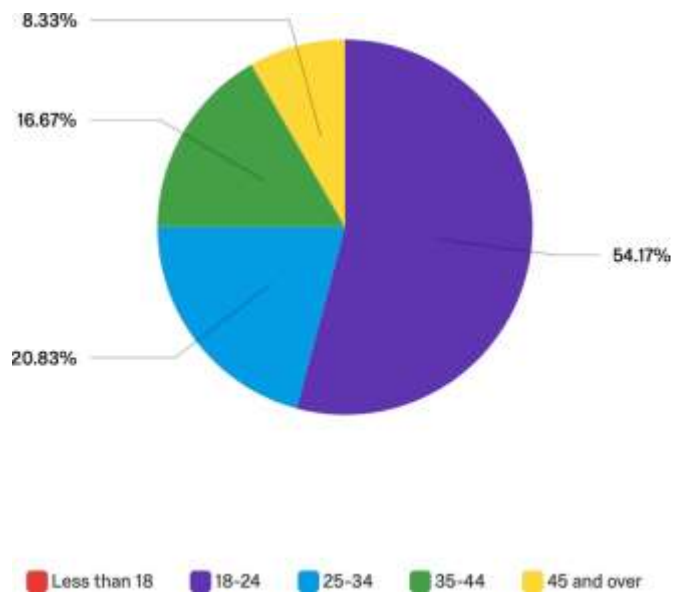


Figure 1 Gender breakdown of participants

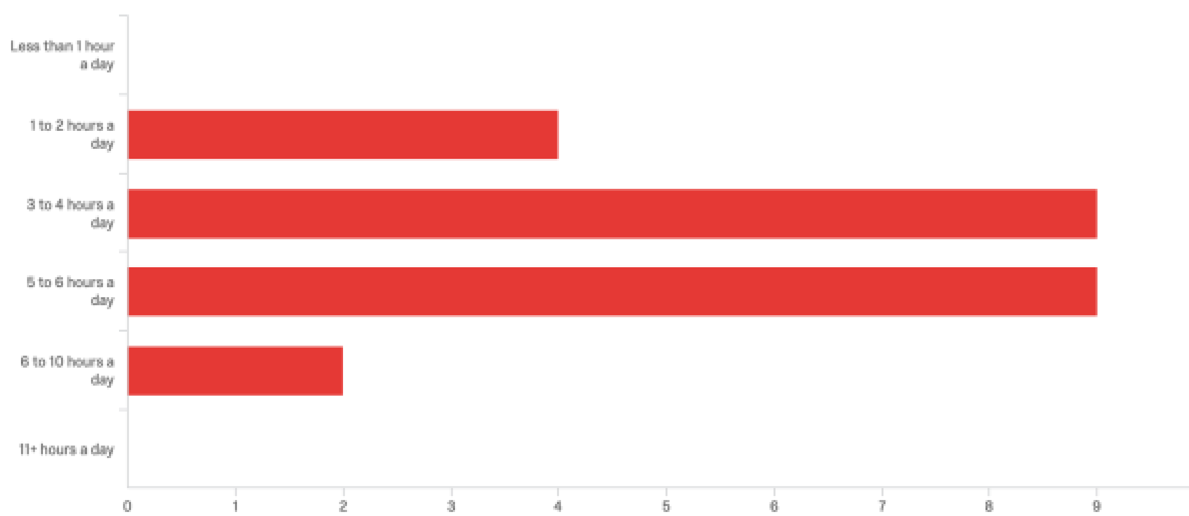


Figure 2 Daily smartphone usage (in hours)

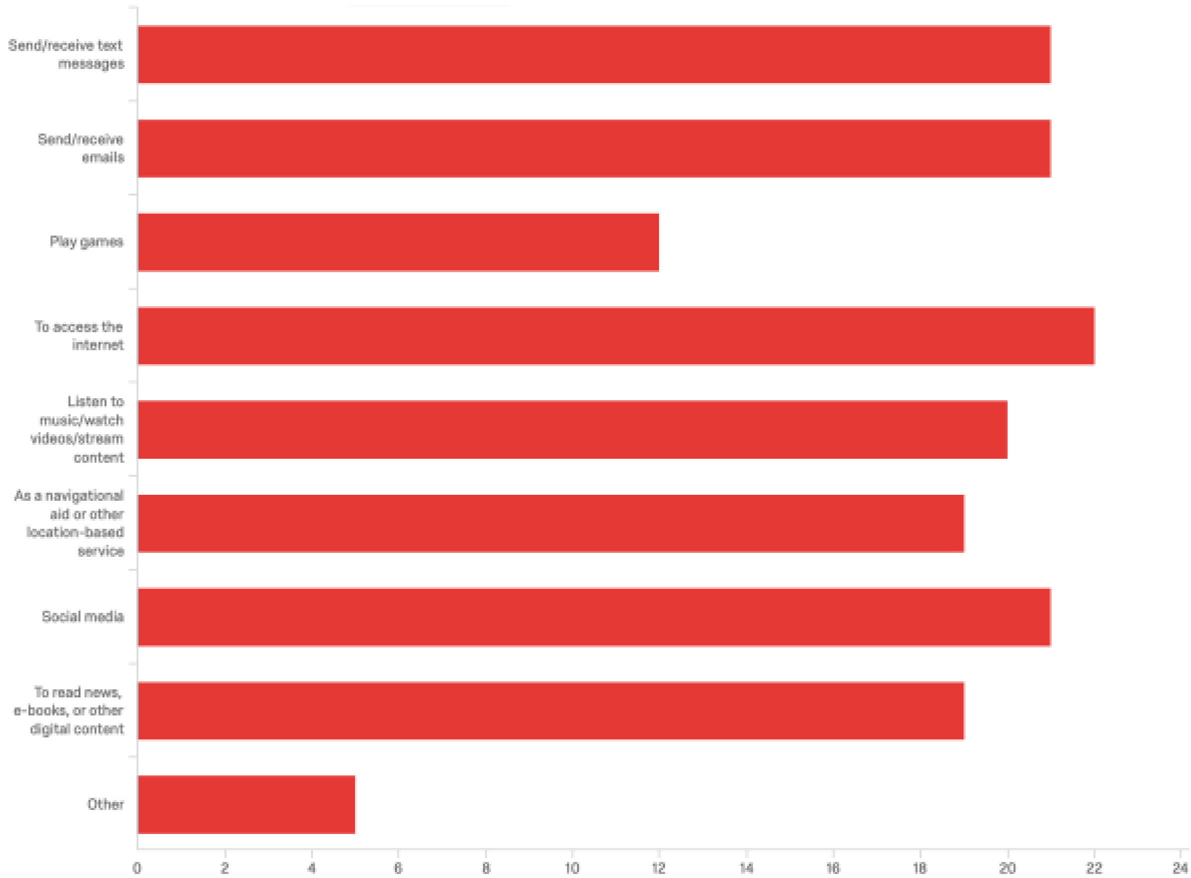


Figure 3 Smartphone usage

Nineteen participants reported their vision being 20/20 either with or without glasses or contacts (with 2 who chose not to answer) and 22 of 24 participants reported that they were not colorblind (with 2 who chose not to answer).

3.2 Materials and Location

The purpose of this study was to measure the effects of performance on a primary cognitive task when introduced with information from an irrelevant secondary task. The information displayed for the secondary task was presented in AR on a Microsoft HoloLens with an Intel 32 bit architecture with TPM 2.0 support, a custom Holographic Processing Unit (HPU 1.0), 2GB RAM, 64GB Flash memory, Windows 10 operating system (OS), weighing 579g (1.27648 lbs) (Microsoft, 2017). The optical hardware consists of two HD 16:9 light engines with a holographic resolution of 2.3M light points and a density of less than 2.5k light points per radian (Microsoft, 2017).

The primary task of identifying the signals was created using Unity and the secondary physical task, the matching game, was created using Unreal Engine. The goal of the game was to match as many numbers on a 10 x 10 grid in four minutes as possible (see Figure 4). These numbers regenerated for the entire four minutes so the participant never ran out of matches. Every three seconds, numbers on the board randomly changed thus making it a constant attention task. The participant's score was the number of correct matches in the allotted four minutes. This game was played on a 16GB RAM computer with a 17 inch 4k Asus monitor within the Games Innovation Lab using a USB corded mouse as the input device.

The experiment took place in an enclosed room within the Games Innovation Lab in KNOY 304 at Purdue University, West Lafayette, Indiana.

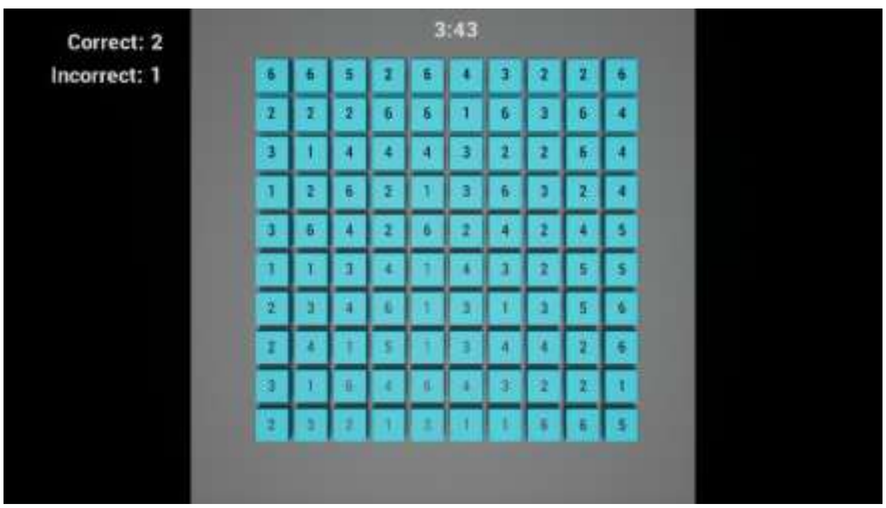


Figure 4 Constant task matching game in Unreal Engine

3.3 Procedure

The experiment consisted of two separate parts which involved separate trials:

1) Having the participants report the content of what they see by either identifying the name of the shape that appears or reading the name of the shape in text format.
2) Having the participants report the presence of either the shape or the name of the shape in text format when it appears by saying "yes."

Since this was a within-subject experiment, participants were exposed to the following three conditions: 1) Condition A (control): HoloLens but with no AR notifications, 2) Condition B: HoloLens with shapes, and 3) Condition C: HoloLens with text. Randomization and counterbalancing was used to reduce conflicting variables. It should be noted that while the control group was not shown any notifications, they were still wearing the HoloLens to remove any inconsistent variables such as visibility of the game, limited field of view, or other physical or ocular restrictions.

Participants were randomly placed in either the content reporting group first or the presence reporting group first. For example, if a participant was placed in the content reporting group first, that person would report the content of the shapes and text as well as perform the control trial, wait five minutes, then report the presence of the shapes and text as well as perform the control trial. The order and number of shapes and text was the same for each trial within the groups (three trials for the content group and three different trials for the presence group minus the control trials where there were no signals) but the participants were asked to report different things in each trial (content and presence). They were then given a five minute break before the next trial began.

Participants were given the task of identifying either a shape or the text name of the shape when it appears in Condition B or C (participants did not know ahead of time the condition they were experiencing). The shapes shown were: triangle, square, circle, star, and diamond. The text was simply the word of the shapes previously mentioned (see Figure 5). The shapes and text were white (RGB 255, 255, 255) per the Microsoft HoloLens design guidelines (Microsoft, 2018b). In addition to color, the font of the text was Arial, which was used in Mustonen's experiment (Mustonen et al., 2013). During the practice round, all participants were shown each shape to ensure each could be correctly identified.

It was stressed that completion of the primary task took precedence over the secondary task but the participants must try to do both to the best of their abilities. The shape/text appeared in the upper right hand corner of the HoloLens as opposed to the center so as to not occlude the primary task of playing the matching game. The right hand side was also selected because both Mac and Windows operating systems have their notification centers in this location (Apple Inc., 2018; Microsoft, 2018a) so expectations of participants matched this design. The shape/text (referred to in the following context as "signals") were displayed for two seconds at random intervals generated between 5-20 seconds apart then disappear. The timing of the signals were determined via a custom Python program that randomized the signals and seconds of when they would appear so as to reduce researcher bias.

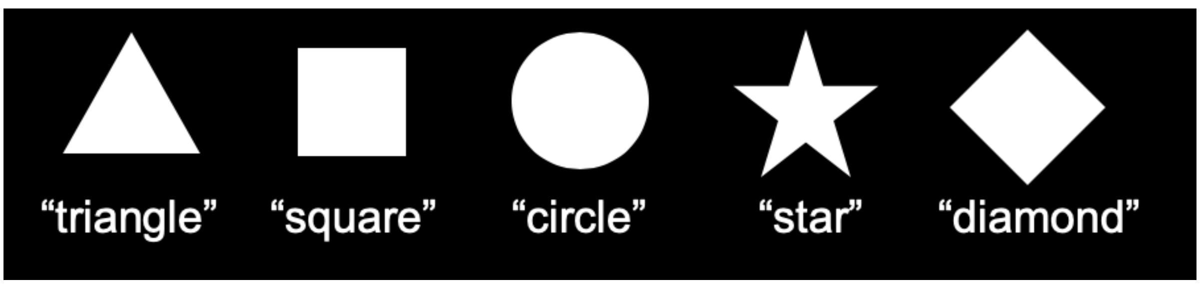


Figure 5 Shape and text signals

Participants were welcomed then seated in front of a table and given the IRB protocol which included a set of instructions explaining the experiment. The researcher was there to answer any questions or to clarify the procedure. Participants were also informed that they could choose to leave the experiment at any time without repercussion. The researcher then asked if there was any questions then had the participants sign the IRB form and fill out the questionnaire.

The audio recorder was tested by the researcher to make certain that the participants' voices were clearly heard. Participants were shown how to put the HoloLens over their eyes and tighten the bands so the device was securely fastened to their heads. They were then shown a sample of the non-task related information so they understand where to look and what the holograms looked like. Participants then had a practice round interacting with the computer game with time for additional questions before the experiment began.

During the actual experiment, participants were shown a 30 second countdown through the HoloLens and counted down out loud once the timer reached the number five. When the timer reached zero, the researcher clapped which marked the starting point for the audio recorder and the participant clicked "Start" on the computer screen and began the matching game which continued for four minutes. If the participant was exposed to Conditions B or C during this time, signals would appear for two seconds then disappear throughout the four minutes of the game. Once the time limit was up, the game ended and exited to a summary screen where the researcher recorded the score.

The audio recordings of the participants' voices were reviewed by the researcher once the experiment was complete. Timing of responses were measured via waveforms where the beginning of the peak was used as the moment of response. In instances where there was no visible waveform or the measured response was too close to being inside or outside the response window, the recording was played back frame by frame to identify the first instance of a phoneme to confirm the accuracy of the response.

The method of measurement was adapted from the walking task experiment done by Mustonen, Berg, Kaistinen, Kawai, and Häkkinen (2013) and the following was measured: four response types (hit, miss, false alarm, and correct rejection). A "hit" was correctly identifying the signal when it is present, a "miss" was when the signal is present but not identified, a "false alarm" was when the signal is identified but is actually not present, and a "correct rejection" was when there is no signal and nothing is identified (Proctor & Van Zhandt, 2008, p. 90). Because the calculations of each measure can be found by using (miss rate = (1 – hit rate)) and (correct rejection
= (1 – false alarms)), only the hit rate and false alarm rate were used to measure the accuracy of the participants' responses to the primary task.

The function of $d' = z(H) - z(F)$ was used to determine reaction times (RT) where H = hit rate, F = false alarms, and *d'* is the dependent variable of response rate of "detectability" or "sensitivity" (the ability of the participant to detect a signal) (Mustonen et al., 2013). Since *d'* is the difference between z-scores and reflects the distance between both the signal and the signal + noise, the function z(H) – z(F) was used instead of H – F. Hits were counted only if the participant responded to the signal (e.g., shape/text) within the two second time limit from the time when the signal first appears. If the participant did not respond within the two second window of the appearance of the signal, it was considered a "miss". However, if the participant responded up to five seconds after the signal disappeared, it was still considered a "miss". This was classified for the purpose of identifying delayed reaction responses and important to note because they were responding to a signal that was indeed present. If the participant responded anytime immediately after already responding to a signal (i.e. a false positive or suggestion of a "double image"), or any time after seven seconds of a signal disappearing but there was no signal actually present, it was considered a "false alarm." Simply put, a response was considered a "false alarm" if it did not fit the criteria for either a "hit" or "miss" (see Figure 6).

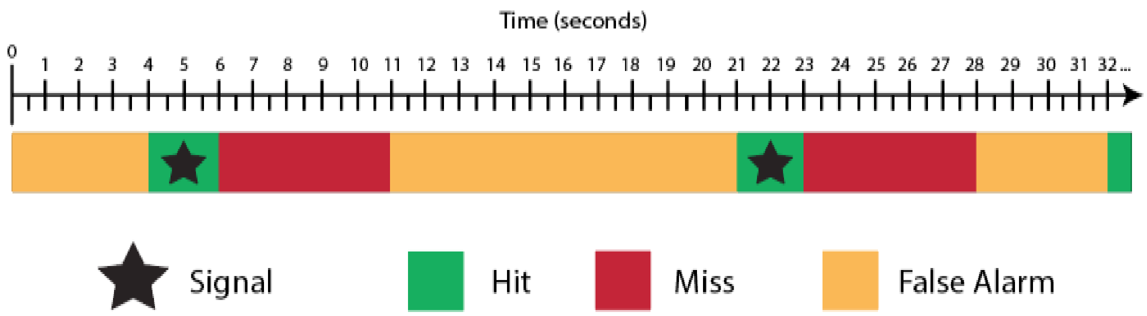


Figure 6 Sample visualization on how "hits", "misses", and "false alarms" were identified

CHAPTER 4. RESULTS

What was learned from this thesis is that it is easier for people to recognize text as opposed to shape in this context of sitting and playing a computer game.

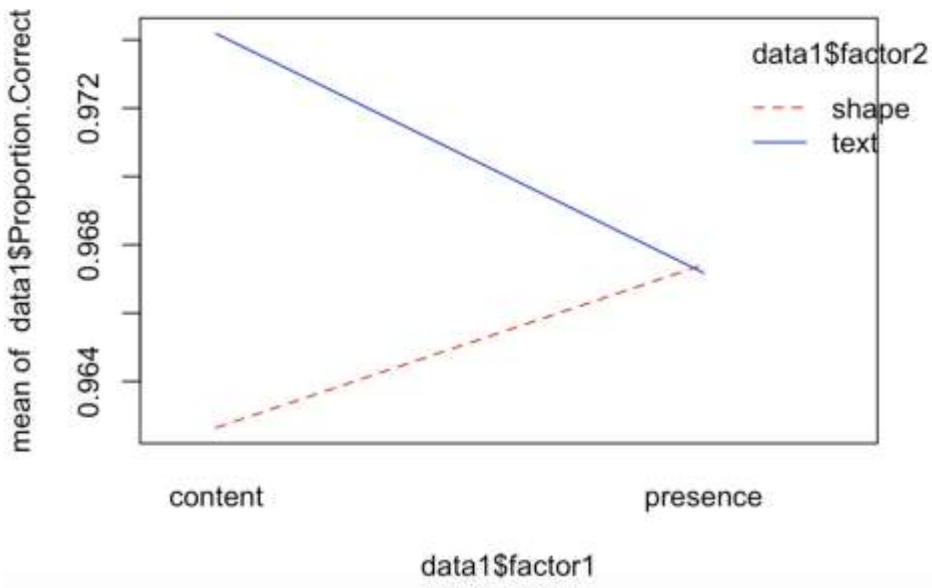
There were three questions that guided this research:

1. Does the format of notification (text or shapes) affect the performance of a non- related primary task and, if so, which affects performance the greatest?
2. When content is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?
3. When presence is of importance, which visual notification (text or shapes) is processed faster while degrading the performance of the primary task the least?

The hypothesis (H_1) states that one of the notifications (text or shape) will have a significantly greater effect than the other on the performance of the primary task while the null hypothesis (H_0) states that neither of the notifications will have an effect on the efficiency on the primary task.

A two-way ANOVA model with random effect across subjects was performed with buildup correlation between the subjects' individual responses and a likelihood ratio test was completed to assess the significance of content/presence and shape/text and their interaction. To adjust for proportion of signals that were equal to zero, the equation 1/(2*N) was used to calculate the z- score (where N = number of possible false alarms during that trial, or the number of possible hits depending on which proportion was being calculated) (Stanislaw, 1999). Similarly, to adjust for proportions with values equal to 1, the equation 1-(1/(2*N)) was used (Stanislaw, 1999).

For correct/incorrect game scores, the chi-squared interaction test was not significant (p = 0.08613) (see Figure 7). Therefore, there is not enough evidence to conclude that there is a significant effect on proportion of correct scores between shape/text and content/presence. The main effect of content (shape/text) on the scores was also not significant (p = 0.1067) meaning that the proportion of correct scores behaved similarly between shape and text. Thus, there is not enough evidence to conclude that shape or text has a significant effect on proportion of correct scores. The main effect of content and presence was also not significant (p = 0.7477) so therefore the proportion of correct scores behaved similarly between content and presence. There is not enough evidence to conclude that content and presence had a significant effect on proportion of correct scores. It is worth mentioning that the control trials were not used in the overall calculations because there was no “hit”, “miss”, and “false alarm” data to record. As such, the scores for the game were not counted as the signal data could not have empty values. However, an average percentage of increase and/or decrease was calculated and revealed an overall 5.71% increase in game scores from the first trial to the last. This indicates a learning



curve effect which fits the results of the game score chi-squared interaction test.

Figure 7 Interaction plot for correct/incorrect game scores

For the proportion of “hit” signals, the chi-squared interaction test was not significant (p = 0.1502) (see Figure 8). There is not enough evidence to conclude that there is a significant effect on proportion of “hit” signals between shape/text and content/presence. When testing the main effect for content (shape/text), the result was significant ($p < 0.001$). Therefore, there is enough evidence to conclude there is a significant effect of text/shape on the proportion of number of “hits”. The main effect of content and presence was also not significant (p = 0.7236) so therefore the proportion of “hit” signals behaved similarly between content and presence.

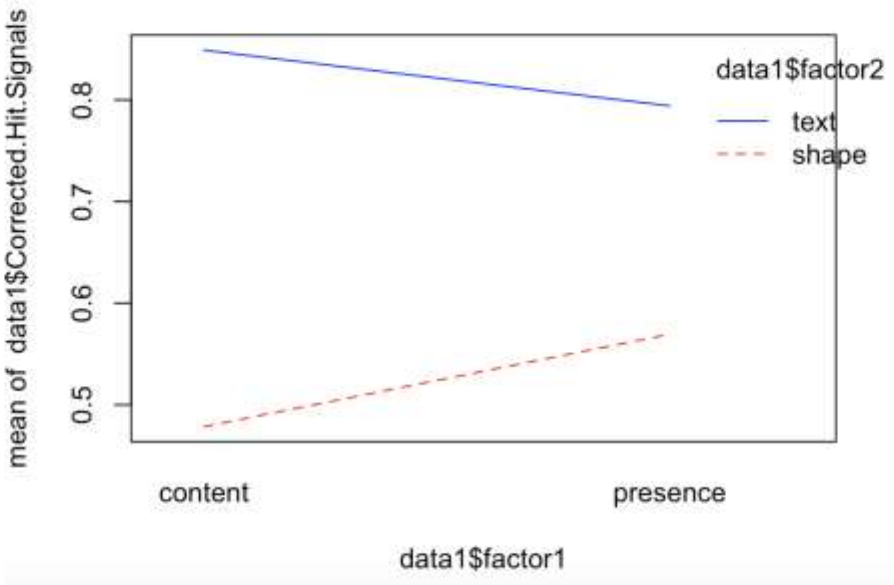


Figure 8 Interaction plot for proportion of "hit" signals

For *d'*, the chi-squared interaction test was not significant ($p = 0.6073$) (see Figure 9). There is not enough evidence to conclude that there is a significant effect on *d'* between shape/text and content/presence. When testing the main effect for content (shape/text), the result was significant ($p < 0.001$). Therefore, there is enough evidence to conclude there is a significant effect of text/shape on the proportion of number of *d'*. After computing the means for text and shape, text had a higher average than shape. The main effect of content and presence was not significant ($p = 0.3567$). There is not enough evidence to conclude that content and presence has a significant effect on proportion of *d'*. A separate calculation of means was done to determine sensitivity levels by trial and signal type (see Table 1). The data shows that participants were more sensitive to the text signal during both content and presence trials. However, it cannot be concluded from this data that text performed better overall which is why the ANOVA was used.

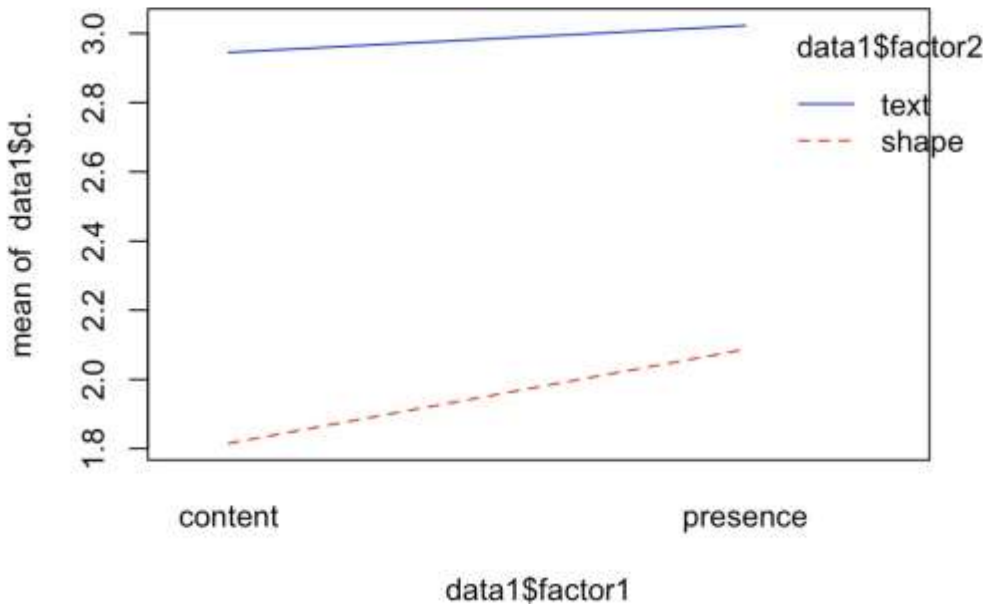


Figure 9 Interaction plot for *d'*

Table 1 Means of *d'*

Type of Signal	Content	Presence
Shape	1.815	2.087
Text	2.975	3.023

In summary, it appears that it is easier for people to recognize text as opposed to shapes in this context of sitting and playing a computer game. While the content or presence was not significant, participants appeared to have a higher sensitivity to text over shapes. When looking at the number of correct matches in the game, it was expected that this number would decrease with either shape or text but the opposite happened. The strange result of the game score actually increasing with both the text and shape signals was attributed to a learning curve wherein the participants got better at the game the more they played. This result is surprising as the literature stated that the performance of the primary task would decrease when a secondary irrelevant task was introduced.

CHAPTER 5. DISCUSSION

5.1 Discussion of Results

The purpose of this study was to determine if there was a difference in the way people responded to two types of signals (text and shape) in augmented reality given two conditions (reporting content or reporting presence). It was hypothesized that one of the signals (text or shape) would have a significantly greater effect than the other on the performance of the primary task.

The results show that the game score (i.e. the primary task) was not affected by either shape/text or content/presence. This can been seen by the previously stated learning curve effect. As the participants continued to play the game, their scores actually increased. Only one participant did worse over time but the mean was very small ($\propto$ = -0.023). This seems to contradict previous research, however response times were not calculated in this experiment due to lack of precision of the instruments used (He et al., 2015; Mustonen et al., 2013; Woodham et al., 2016).

While the proportion of “hit” signals behaved similarly between content and presence, there seems to be a significant effect of text/shape on the proportion of number of “hits”. This is also shown in the *d’* of participants who appeared to have a higher sensitivity for text signals during both content and presence trials. This could be due to the difficulty and increased time it takes one to discern what shape one is looking at (i.e. counting sides and/or holistically processing the shape itself). Location could also have been a factor; the center point of the text and shape were kept the same but the text takes up more space horizontally and could have been more intruding in one’s field of view, making it more obvious to see. This contradicts the suggestion by Sabelman and Lam (2015) who stated that that minimalistic symbols and little to no text was ideal for pilots.

These results taken as a whole suggest that under this particular condition of sitting and playing a constant task game, one type of signal appears to be better than another. However, the literature review strongly argues that the best way to present information will depend on the condition and task being performed. Because a learning curve was found in this study, future best practices for augmented reality wearables should perhaps include a training period to accommodate users to the new addition to their vision. It would be interesting to build upon this information by performing vigilance tasks with shape/text signals in different everyday scenarios. It might also be beneficial to analyze different sizes, locations, and spatial arrangements of different signals to not only to see if the learning curve can be replicated but to also find an optimal layout to limit distractions.

5.2 Limitations and Future Directions

There were several limitations of this study, many involving the implementation of the experiment itself, that should be considered when looking at the results. First, the timing between the two systems (HoloLens and computerized matching game) were not linked or synchronized in any way, resulting in fluctuations of participant response time and correlated "hits" and "false alarms". Some participants were even slow to hit the "start" button on the game which threw off their response time. The data was not adjusted for delayed response times, resulting in some outliers where it appeared that the participant missed most of the signals. Second, audio recordings of participant responses had to be analyzed subjectively when there were no visible waveforms to use as markers. Even when there were visible waveforms, most of the recordings had to be re- analyzed for audible phonemes to get an accurate response time. Third, there were problems with wearing the HoloLens itself; the device slid down several times and participants had a hard time seeing any and sometimes all the signals. Some users reported seeing the countdown timer but then did not report any signals while other reported the first half of the set of signals and missed the last half. The angle of the HoloLens was also an issue as one participant reported the signals "blinking" which was due to the angle of the person's head and the angle of the lenses. Fourth, some participants misidentified the signals (e.g., said "circle" instead of "yes" during presence trials) which were marked as a "miss." Even after giving the participants the instructions, some even identified the text signals as "word… diamond" until they were corrected.

Various behavioral observations were made throughout the study which included participants constantly adjusting the HoloLens, tilting their heads up to get a clearer view of the computer screen, and, muttering words or phrases in frustration. One participant reported that "Just knowing that something *might* pop up is distracting."

This study provides foundational data for displaying information on an untethered augmented reality headset. As observed in the literature review, a study like this has not been done before and the methodology was borrowed heavily from existing signal theory and human factors studies. There is potential to explore further and develop best practices for displaying information in augmented reality. It would be ideal to obtain a larger sample size and even observe the differences between age groups and people of varying abilities to make future wearable augmented displays as accessible as possible.

REFERENCES

Apple Inc. (2018). Use notifications on your Mac. Retrieved April 7, 2018, from https://support.apple.com/en-us/HT204079

Ayers, J. W., Leas, E. C., Dredze, M., Allem, J. P., Grabowski, J. G., & Hill, L. (2016). Pokémon GO-A new distraction for drivers and pedestrians. *JAMA Internal Medicine*, *176*(12), 1865– 1866. https://doi.org/10.1001/jamainternmed.2016.6274

Bark, K., Tran, C., Fujimura, K., & Ng-Thow-Hing, V. (2014). Personal Navi: Benefits of an augmented reality navigational aid using a see-thru 3D volumetric HUD. *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 1–8. https://doi.org/10.1145/2667317.2667329

Cassagnol, D. (2018). AR/VR will open eyes in 2018 : Record-setting U.S. sales projected, says CTA. Retrieved February 24, 2018, from https://www.cta.tech/News/Press-Releases/2018/January/AR-VR-Will-Open-Eyes-in-2018-Record-Setting-U-S.aspx

Cook, J. L., & Jones, R. M. (2011). Texting and accessing the web while driving: traffic citations and crashes among young adult drivers. *Traffic Injury Prevention*, 12(6), 545- 549. https://doi.org/http://dx.doi.org/10.1080/15389588.2011.620999

Haiduk, P. M. (2017). A flight guidance display format on smart glasses for private pilots. *Aviation Psychology and Applied Human Factors*, *7*(2), 67–78. https://doi.org/http://dx.doi.org/10.1027/2192-0923/a000119

He, J., Choi, W., McCarley, J. S., Chaparro, B. S., & Wang, C. (2015). Texting while driving using Google Glass: Promising but not distraction-free. *Accident Analysis and Prevention*, *81*, 218– 229. https://doi.org/10.1016/j.aap.2015.03.033

Hou, L., Wang, X., Bernold, L., & Love, P. E. D. (2013). Using animated augmented reality to cognitively guide assembly. *Journal of Computing in Civil Engineering*, *27*(5), 439–452. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000184.

Joseph, B., & Armstrong, D. G. (2016). Potential perils of peri-Pokémon perambulation: The dark reality of augmented reality? *Oxford Medical Case Reports*, *0*(10), 265–266. https://doi.org/10.1093/omcr/omw080

Kipman, A. (2016). Announcing Microsoft HoloLens Development Edition open for pre-order, shipping March 30. Retrieved February 20, 2018, from https://blogs.windows.com/devices/2016/02/29/announcing-microsoft-HoloLens-development-edition-open-for-pre-order-shipping-march-30/#uU5MecD8EfIDYLtH.97

Martínez, H., Skournetou, D., Hyppölä, J., Laukkanen, S., & Heikkilä, A. (2014). Drivers and bottlenecks in the adoption of augmented reality applications. *Journal of Multimedia Theory and Applications*, *2*, 27–44. https://doi.org/10.11159/jmta.2014.004

Mayer, R. E., & Moreno, R. (1998). A split-attention effect in multimedia learning: Evidence for dual processing systems in working memory. *Journal of Educational Psychology*, *90*(2), 312– 321. https://doi.org/10.1037/0022-0663.90.2.312

Merriam-Webster Incorporated. (2018). Augmented reality. Retrieved February 3, 2018, from https://www.merriam-webster.com/dictionary/augmented reality

Microsoft. (2017). HoloLens hardware details. Retrieved March 4, 2018, from https://developer.microsoft.com/en-us/windows/mixed-reality/HoloLens_hardware_details

Microsoft. (2018a). Find your notifications. Retrieved April 7, 2018, from https://support.microsoft.com/en-us/help/17203/windows-10-find-notifications

Microsoft. (2018b). Typography. Retrieved from https://docs.microsoft.com/en-us/windows/mixed-reality/typography

Mustonen, T., Berg, M., Kaistinen, J., Kawai, T., & Häkkinen, J. (2013). Visual task performance using a monocular see-through head-mounted display (HMD) while walking. *Journal of Experimental Psychology: Applied*, *19*(4), 333–344. https://doi.org/10.1037/a0034635

Norman, D. A., & Bobrow, D. G. (1976). On the analysis of performance operating characteristics. *Psychological Review*, *83*(6), 508–510. https://doi.org/10.1037/0033-295X.83.6.508

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*(2), 220–244. https://doi.org/10.1037/0033-2909.116.2.220

Phelan, D. (2019). Apple AR glasses to launch late 2019 or 2020. Retrieved April 13, 2019, from https://www.forbes.com/sites/davidphelan/2019/03/08/apple-ar-glasses-to-launch-late-2019- or-2020-analyst-says/#5081bbc1613e

Proctor, R. W., & Van Zhandt, T. (2008). *Human Factors in Simple and Complex Systems* (Second ed.). Boca Raton: CRC Press.

Rane, P., Kim, H., Marcano, J. L., & Gabbard, J. L. (2016). Virtual road signs: Augmented reality driving aid for novice drivers. *Proceedings of the Human Factors and Ergonomics Society*, *60*(1), 1748–1752. https://doi.org/10.1177/1541931213601401

Rummer, R., Schweppe, J., Fürstenberg, A., Scheiter, K., & Zindler, A. (2011). The perceptual basis of the modality effect in multimedia learning. *Journal of Experimental Psychology: Applied*, *17*(2), 159–174. https://doi.org/10.1037/a0023588

Sabelman, E. E., & Lam, R. (2015). The real-life dangers of augmented reality. *IEEE Spectrum*, *52*(7), 48–53. https://doi.org/10.1109/MSPEC.2015.7131695

Santos, C., Brunelli, M., Araujo, T., Carneiro, N., Marques, A., Mota, M., … Meiguins, B. (2016). Guidelines for graphical user interface design in mobile augmented reality applications. In *Lecture Notes in Computer Science* (Vol. 9470, pp. 71–81).

Sawyer, B. D., Finomore, V. S., Calvo, A. A., & Hancock, P. A. (2014). Google Glass: A driver distraction cause or cure? *Human Factors: The Journal of Human Factors and Ergonomics Society*, *56*(7), 1307–1321. https://doi.org/10.1177/0018720814555723

Stanislaw, H. (1999). Calculation of signal detection theory measures - art%3A10.3758%2FBF03207704.pdf. *Behavior REsearch Methods, Instruments & Computers*, *31*(1), 137–149. https://doi.org/https://doi.org/10.3758/BF03207704

Vogel, J. R., Schultz, M. C., & Schultz, J. T. (2004). The effect of human factors on the helmet-mounted display. https://doi.org/10.1016/j.ecns.2012.07.047

Whittinghill, D., & Herring, D. G. (2017). A comparison of text-annotated vs. purely visio-spatial instructions for video game input sequences. *Computers in Entertainment*, *15*(2), 1–19. https://doi.org/10.1145/2700527

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of Human Factors and Ergonomic Society*, *50*(3), 449–455. https://doi.org/10.1518/001872008X288394.

Woodham, A., Billinghurst, M., & Helton, W. S. (2016). Climbing with a head-mounted display: Dual-task costs. *Human Factors: The Journal of Human Factors and Ergonomics Society*, *58*(3), 452–461. https://doi.org/10.1177/0018720815623431

Yeh, M., Merlo, J. L., Wickens, C. D., & Brandenburg, D. L. (2003). Head up versus head down: The costs of imprecision, unreliability, and visual clutter on cue effectiveness for display signaling. *Human Factors: The Journal of Human Factors and Ergonomics Society*, *45*(3), 390–407. https://doi.org/10.1518/hfes.45.3.390.27249