Augmented Virtuality for Head-Mounted Displays

Improving User Experience by Incorporating the Real into the Virtual

Jacob H. Clarkson
University of Cape Town
Rondebosch
Cape Town, South Africa
jacobhclarkson@gmail.com

1. INTRODUCTION

Consumer level Virtual Reality (VR) experiences are about to arrive in the form of Head-Mounted Displays (HMDs) such as the Oculus Rift [7] and HTC Vive [5]. VR has seen a massive resurgence in recent years, and by 2020, the industry is expected to be worth approximately $30 billion [4]. With about half of the market being accounted for by VR games [4], computer gamers are the key demographic for consumer HMDs [16]. Although projections like this may seem overly optimistic, technology has come a long way since the 90s VR bubble burst. The problems previously faced by consumer grade VR experiences (such as low resolution, small field of view, uncomfortable weight, motion-sickness inducing latency, and poor graphical quality [13]) have been largely addressed [16].

Given the expected ubiquity of HMDs, it is important to consider how and where they will be used, and who will be using them. Since personal entertainment (video games in particular) will be a primary function of consumer level VR, it is of particular importance to consider the context of home usage [16]. Despite the fact that video games are meant to place players in a virtual world, to the potential (if not deliberate) exclusion of the real world, a considerable number of real world interactions still take place. Examples of these interactions include the use of peripherals such as a keyboard, or taking a drink. The complete visual exclusion of the real world by HMDs, while affording greater immersion, makes interactions as trivial and natural as these far more difficult.

For VR gaming to be as fully embraced as traditional gaming (using conventional computer monitors), real world interactions commonly performed by gamers should be facilitated. Augmented Virtuality may offer a solution. Augmented Virtuality (AV), as opposed to Augmented Reality (AR), is the addition of elements of the real into the virtual [19]. A forward facing camera affixed to the front of an HMD could give the wearer a view of reality, effectively letting the user "see through" the device. Allowing users to see the real world through their HMD would facilitate real world interaction without forcing the user to completely disengage from what they are doing by removing it.

The aim of this research is to modify an existing approach to AV and evaluate its effectiveness in terms of facilitating real word interaction, and preserving the sense of immersion experienced by the user.

2. RELATED WORK

VR attempts to heighten the way that people experience virtual environments. Technologically, the largest difference between VR and a conventional monitor setup is the display. Incarnations of VR such as CAVE, and "wedge" [13], for example, project multiple views of the virtual environment onto the walls of a room in which the user stands, whereas HMDs use two small screens, one in front of each eye, to create the illusion of depth [21]. Many implementations of VR also make use of positional tracking, such that the movements of the user are directly translated into movements in the virtual environment. Room scale VR utilizes this positional tracking to a larger extent than HMDs (with the possible exception of the HTC Vive), allowing users to walk around freely in the real/virtual room. In contrast, the Oculus Rift tracks only the user's head, and is intended for more of a "sitting down" experience with more conventional controls. The more affordable and practical nature of HMDs, as opposed to room scale VR, is the reason they are the main focus of this research. Despite the various differences between VR implementations, they share a common goal: a heightened sense of Presence [24].

Presence [22] can be loosely defined as the user's sense of "being" in a virtual environment. For an individual experiencing Presence in a virtual environment, equipment (displays and controllers) and physical surroundings fade away, leaving the impression that the virtual world that they are in is more real than their actual reality at that point. In short, Presence is the feeling of being in one place, while physically being in another [28]. VR strives to create this experience to a far greater extent than is possible using normal computer monitors. Presence is very important for games, and the effectiveness of virtual environments has often been linked to the amount of Presence experienced by users [28]. As noted by Weibel and Wissmath [27], the pleasure of being immersed in a virtual world is a primary reason to play games, and Presence is a core component of immersion.

Of the various ways in which VR can be implemented, the HMD is perhaps the most practical for consumers. HMDs are small, light, and relatively inexpensive, while offering high fidelity visuals and audio. However, there is one major problem with the way they create a sense of Presence: they cut the wearer off visually from their surroundings. Not only are users cut off from the tools and peripherals they may need for whatever activity they are engaged in (such as input devices) [8], but also from all other real-world objects they may wish to interact with. The argument could be made that the entire point of video games, immersive or not, is to cut the user off from the real world. Why then should it matter, if users are now visually cut off from that world as well? The answer is manifold.

2.1 Problems with HMDs

Certain tasks, such as text entry, are very difficult to accomplish without the sense of sight [16]. Although a subset of keyboard commands, or the mapping of a controller, can be memorized and used for gameplay [8], these are inadequate for tasks that require a greater bandwidth of input. In fact, current HMD users rate the inability to provide input into VR as as greater impediment to widespread adoption than the hitherto rather common HMD induced nausea [16].

Text-entry, especially in online multiplayer games, is remarkably prevalent. Almost every online multiplayer game has a text chat box in a corner of the heads-up display [20]. An analysis of communication in the game *World of Warcraft* [1] showed that nearly half a million text messages were sent in the 11775 recorded sessions, an average of about 40 text messages per session (where the average session lasted 57 minutes) [25]. This amounts to an average of more than one text message being sent every two minutes per player. In fact, the dominant form of communication in games of several genres is text [12]. The effectiveness of voice-based communication for cooperative endeavours and task coordination [14] lends itself to use in those situations, but text-based chat is used for socio-emotional communication to a large extent [20, 12]. Text-based communication is seemingly entrenched in the social and emotional aspects of multiplayer gaming. This is perhaps no more evident than in the recent case where text-based communication was not included in the popular *Metal Gear Solid V:MGS Online* [2], leading to outrage amongst the franchise's fan base. To quote one Mr. GGscrub, "Wtf like, it was an essential part of the charm of MGO2! How the hell is it missing?" [3]. The sentiment is echoed by CanOpener74 ("really not a good change, this... sadness... pervades.") and many others [3, 6]. Text-based chat in games is just one example of a very likely form of input required of users that is difficult to perform while wearing an HMD.

The objects that people wish to interact with while they are playing games are not limited to peripherals used for playing the games. Studies have shown that 67% of the time spent playing computer games is also spent doing something else [10]. 40% of the actions constituting that "something else" are real-world interactions not directly linked to the game being played, such as eating or drinking, reading, homework, and using a cellphone [10]. Currently, HMD users rate their ability to interact with objects as extremely ineffective [16].

Text-based communication (and input in general), as well as real-world interactions, are common enough tasks among computer gamers to be given some attention in terms of how we can help users execute them.

2.2 Living with an HMD

Although the first consumer level VR hardware has only recently begun shipping, earlier development versions of the hardware (such as the Oculus Rift DK1 and DK2) have been in use by early adopters for some time. This has allowed for the usage of these devices to be studied, and to find out how people are dealing with the visual cutoff created by HMDs. The "solutions" (if indeed we can call them that) that users have come up with are not particularly good. Chief among them are "peeping", "groping", and isolation. Peeping involves temporarily lifting the device off one's eyes in order to gain awareness of, or interact with, the real world. Groping is an attempt to interact with the real world without removing the HMD. Isolation, a very common behaviour among early HMD adopters (approximately 80% of whom report doing this [16]), involves play sessions taking place with no other people or real-world objects around, a total exclusion of reality. HMD users agree that having to resort to peeping is frustrating [16]. Furthermore, peeping defeats the primary purpose of the HMD, as lifting it off clearly constitutes a significant break in Presence. Isolation, deliberately forgoing all the activities one may usually partake in while playing a game, is unlikely to be accepted by the majority of users, especially considering how prevalent real-world interactions are while playing games. It is important to note here that, apart from peeping, the visual cutoff created by wearing an HMD does not necessarily interfere with the usability of the HMD itself, but rather with the usability of peripherals and other objects. If VR gaming is to be as ubiquitous as traditional gaming is today, the usability issues created by HMDs need to be better addressed.

2.3 Mixed Reality as a Potential Solution

Mixed Reality displays are a subset of VR technologies that merge the real and virtual worlds [18]. All mixed reality displays can be thought of as occupying a point somewhere along a "virtuality continuum" as illustrated in Figure 1. The widely known AR is closer to the Reality side of the spectrum, where the display of an entirely real environment has virtual artefacts added to it (predominantly real with few virtual elements). In AV, on the other hand, certain real objects are made visible in an otherwise entirely virtual environment (predominantly virtual with few real elements). One way in which AV can be implemented is by attaching a digital video camera (such as a webcam) to the front of an HMD to allow the wearer to "see through" the device [18]. By doing so, varying degrees of the real world, as seen by the camera, can be relayed to the HMD and superimposed over the display of the virtual world.

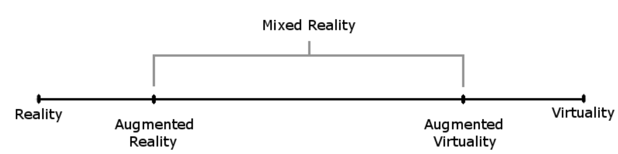


Figure 1: The Virtuality Continuum

A recent study assessed several AV techniques [16]. A series of experiments was conducted with the aim of assessing user performance and preference with respect to typing and real-world interaction in several AV conditions. The scores achieved were compared to a pure virtuality baseline (wearing an HMD and not ever taking it off), and a "peeping" baseline (lifting the HMD off the eyes when required). AV was implemented, in all cases, by passing various amounts of video from a standard webcam attached to the front of the HMD to the display when required. The AV conditions that were tested included full blending, partial blending, and minimal blending. These conditions are depicted in Figure 2.

In the full blending condition (Figure 2a), the display of the HMD changes between showing a full view of the virtual world, and a full view of the real world.

In partial blending (Figure 2b), rather than the full view of reality being displayed, only the user's hands, peripherals, and objects to interact with were inserted into the view of the virtual world.

In minimal blending (Figure 2c), only a small area of the real world around the users hands was inserted into the virtual world.

The methods differ in how much of the real world is shown, with full blending showing the most, minimal blending showing the least, and partial blending falling between the two. It was found that both minimal and partial blending performed better in terms of preserving the user's sense of Presence than when users "peeped". Furthermore, users far preferred partial and minimal blending to the baseline conditions in terms of how easy it was to interact with real-world objects. Finally, partial blending was found to increase typing performance compared to the baseline conditions (typing performance for the minimal blending condition was not assessed).

The research conducted by McGill et al. [16] revealed several useful results: Input and real-world interaction can be facilitated using AV; Only a small amount of reality needs to be displayed for AV to be effective; AV does not have an overly negative impact on Presence. However, there are several issues with the study. The effect of minimal blending on typing performance, compared to virtuality and reality baselines, was not assessed, and despite the insistence that the context of home and office usage be taken into account, both the minimal and partial blending techniques that were implemented are dependent on chroma-keying. This requires a user's entire desk/cubicle to be "green-screened" in order for these implementations to work. This is patently impractical for the average consumer, the target market for HMDs, and the very people whose experience AV is aimed at improving.

A similar study investigated several AV techniques purely in terms of facilitating real-world interaction [9]. AV implementations very similar to minimal, partial, and full blending were assessed, as well as inset AV. Inset AV, proposed as early as twenty years ago, involves using the video pass through technique to create a small inset of reality as an overlay which is superimposed onto the virtual scene [17]. A major benefit of the inset method is that it does not require the use of chroma-keying. Unfortunately, Budhiraja et al. [9] found that inset AV did not perform well compared to their versions of minimal and partial blending, with users reporting difficultly mainly with the size discrepancy between what was displayed in the inset and what the actual sizes of objects were in the real world. Their results with respect to minimal and partial blending conditions were very similar to those obtained by McGill et al. [16], showing good performance in terms of real-world interaction and strong user preference. The minimal blending AV system implemented by Budhiraja et al. relies on color segmentation as opposed to chroma-keying to selectively display aspects of the real world (such as a users hands). While perhaps more practical than chroma-keying, color segmentation can be sensitive to skin tone, lighting conditions, and background contrast [26], all of which would inevitably lead to issues for consumers. A major difference between these two studies is the design of the transitional interface, or the conditions under which reality is shown. Budhiraja et al. [9] allowed the user to dictate, with the press of a button, when the AV system would become active. McGill et al. [16] found that allowing the users to choose when to activate AV scored far lower in terms of Presence and preference than what they call "inferred blending", where the AV system automatically activates based on context (e.g. when the user stretches their hands out in front of them). Budhiraja et al. mention such an augmentation to their system as potential future work.

Minimal blending AV seems to have the most potential for assisting HMD users in the home context. Minimal blending has a lower impact on Presence compared to other AV implementations, and results indicate that it facilitates both input tasks and other real-world interactions well. Furthermore, minimal blending is not strictly dependent on impractical or unreliable techniques such as chroma-keying or colour segmentation. For these reasons, this research will explore minimal blending further, attempting to replicate the positive results of previous work [16, 9] with a new implementation that is more practical for consumers. Finally, this research will measure the impact that AV has on typing performance, an area that has not yet been explored.

3. RESEARCH QUESTIONS

The aim of this research is to investigate the user experience of a version of minimal blending AV. It will be assessed in terms of its impact on Presence, and how well it enables users to complete real world tasks.

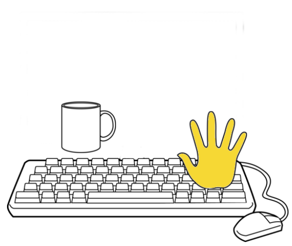
This research is important for several reasons. The fact that users of HMDs cannot see while wearing the device presents several usability issues. This, coupled with the expected popularity of these devices, dictates that we attempt to address these issues while preserving their main function (heightened immersion). In so doing, one of the few remaining obstacles that may prevent widespread adoption of HMDs will hopefully be removed. Previous work regarding minimal blending has shown it to have great promise. However, these implementations have relied on a chroma-key approach, which is clearly not practical for end users. This research will attempt to ascertain whether minimal blending can be as effective without the need to green-screen one's entertainment area.

The primary research question that will be investigated is as follows: *Can minimal blending AV be successfully implemented without the use of chroma-keying?*

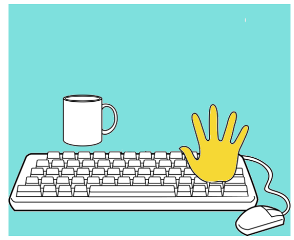
For the purposes of this research, the success of an AV technique is defined by its effect on typing performance, real-world interaction, and Presence. In order, therefore, to answer the primary research question, the following research questions and hypotheses will be investigated:

3.1 Typing Performance

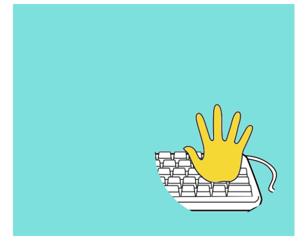
One of the areas this research aims to explore is the viability of AV as a solution to the input problem created by HMD-induced visual cutoff. Of the various examples of input that games require of users, text entry is one of the most challenging, and is required often in the form of text-based communication. Text input is easily and comprehensively measurable, and serves as a difficult example of input in general. The first research question, therefore, is as follows:



(a) Full Blending AV - When the user's hand is detected, the entire display switches to the real world. None of the virtual world is shown.



(b) Partial Blending AV - When the user's hand is detected, the user's hand and any interactive objects are displayed "over" the virtual world. All other parts of the virtual world are still visible.



(c) Minimal Blending AV - Only the user's hand and a small area of the real world around it shown when the hand is detected. The display of the virtual environment is left largely in tact.

Figure 2: Graphical Representation of Different Types of AV - Adapted from McGill *et al.* [16]. The black and white parts of the images represent the real world in front of the user. The blue areas represent areas of the visual field that are being occluded by the display of a virtual environment.

Is typing performance better in AV than in virtuality?

H_{A0} There is no difference in typing performance between AV and a virtuality baseline.

H_{A1} Typing performance is better in AV than in a virtuality baseline.

H_{A2} Typing performance is worse in AV than in a virtuality baseline.

3.2 Real-World Interaction

For people to be able to play VR games in a way that is similar to they way they play conventional games, occasional interactions with real-world objects should be easy. The ease with which a user is able to execute actions, such as taking a drink, can be measured both by the amount of time that such interactions take, as well as the cognitive load that executing this interaction (coupled with what the user is doing in the game) places on the user.

Is it easier to interact with real-world objects in AV than it is in virtuality?

H_{B0} It is as difficult to perform real-world interactions in AV as it is in a virtuality baseline.

H_{B1} It is easier to perform real-world interactions in AV than it is in a virtuality baseline.

H_{B2} It is more difficult to perform real-world interactions in AV than it is in a virtuality baseline.

3.3 Presence

Presence is a crucial part of the gameplay experience, and a primary goal of HMDs. Any attempt to improve the experience of using an HMD must preserve Presence as much as possible.

Is Presence higher in AV than it is in virtuality?

H_{C0} There is no difference in Presence scores in AV and in a virtuality baseline.

H_{C1} Presence scores are higher in AV than they are in a virtuality baseline.

H_{C2} Presence scores are lower in AV than they are in a virtuality baseline

4. METHOD

In order to answer the proposed research questions, an experiment will be conducted to ascertain what impact AV has on Presence, typing performance, and real-world interaction. The apparatus required to conduct these experiments will take the form of an immersive 3D game.

To complete the game, the player will have to complete a series of increasingly complicated tasks. Users will have to input text, as well as glean clues by interacting with the real-world around them, in order to complete these tasks. Text entry and real-world interaction will both be incorporated into the game itself.

4.1 System Design

4.1.1 Design Features

The core design feature of the proposed system is an AV implementation based on the work of McGill et.al. [16]. The proposed system, however, will rely on alternative hand tracking methods that do not require the use of the chroma-key technique. The second important feature will be apparatus to measure the speed and accuracy of text input. Both of these systems will be incorporated into an immersive, VR compatible game, which will be created with the goal of promoting a sense of Presence in the player[1].

4.1.2 Development Platform

The development platform will be the Windows version of Unity3D.

4.1.3 Implementation Strategy

There are three main components to the system: the game; the AV system; and apparatus for text entry analysis. Unity3D allows for a high degree of modularity, so all three components can be developed independently. Since the focus of this research is on AV and its effects, the AV and text entry systems will be developed first. Once these

[1]Ideally, an existing game would be used. Game creation is not necessarily the focus of this research, and brings with it many obstacles and resource requirements. However, no suitable game - that integrates both typing and real-world interaction into the gameplay experience - currently exists. Such a game will be created. This hurdle is not insurmountable, as the researcher has experience developing games in Unity3D.

two systems are functional, the game itself will become the major focus.

4.1.4 Expected Challenges

The major challenge will be the creation of the AV system. This system needs to operate with low latency, and high fidelity, in order to yield the best results. Furthermore, the smooth operation of this system requires that three separate pieces of technology be made to work seamlessly together. Namely, an Oculus Rift, a Leap Motion controller, and a webcam.

4.2 Experimental Design

The experiment conducted will be a single-blind, between-groups design. The participants will be randomly allocated to one of two groups, and will not know which group they are in, or what the nature of the other group is. Participants will not be informed what the experiment aims to measure, but merely that they are to play a game and answer some questions about their experience afterwards.

The two groups are as follows:

1. Control Condition (Virtuality)
2. Experimental Condition (Augmented Virtuality)

In both cases, participants will play a game while wearing an HMD. During the course of gameplay, they will enter text using a keyboard, and interact with several real-world objects. In the Control Condition, all elements of the AV system will be disabled. The AV system will be fully active in the Experimental Condition.

A third group is not necessary, as baseline typing and real-world interaction speed will be captured for each participant post-experiment[2]. A between-groups design was chosen to limit contamination by extraneous factors, particularly learning effects and fatigue. However, this choice of design will require a larger number of participants in order for any effects to be detected to a reasonable degree of significance (compared to a within-subjects design).

The independent variable (IV) in this case is group (defined by the absence or presence of the AV system). The dependent variables (DVs) are Presence, ease of real-world interaction, and typing performance.

4.3 Procedure

4.3.1 Participants

Pseudo-random convenience sampling will be used to recruit participants. It is highly likely that none of the participants that are recruited will ever have had experience with immersive VR. As such, participants with computer and computer gaming experience will be favored. Participants with little to no computer/gaming experience will likely be overwhelmed by the plethora of new and unfamiliar technology in the experiment, which may affect their performance and responses. Computer Science undergraduate and Honours students will be the easiest to access, and likely to meet the stipulated criterion. People with severe visual impairments, or whose ability to perceive in 3D is in any way compromised, will be excluded. Because time to first key press is one of the metrics involved in text entry analysis, touch typists will not be excluded from the study, as this is one area that may still be heavily affected by the absence or presence of the AV system. Whether or not candidates are able to touch type will be recorded nevertheless, in case their inclusion results in otherwise inexplicable outliers in the gathered data.

4.3.2 Pre-Experiment

Before the experiment begins, participants will be given a handout describing the experiment and the technology used therein. Care will be taken not to allude to any of the measures being used, in order to prevent bias. The use of a handout, rather than verbal instructions, will ensure a consistent pre-experiment experience for all participants, and reduce the risk of experimenter bias. Furthermore, non-verbal instructions will serve to discourage potentially contaminating dialogue with the experimenter. The handout will serve as informed consent, and participants will sign if they are willing to proceed.

4.3.3 Experiment

During the course of the experiment, participants will play the game while wearing an HMD. An in-game "instructor" will guide the player through the game by way of text prompts. These prompts will either give players text "codes" that need to be entered in order to proceed, or instruct the participant to interact with a specific object in the real world that will give them a clue as to how to proceed. There will be a total of six tasks that need to be completed in the game. These tasks will take place in different locations throughout the world, and participants will have time between completing each task to walk around the virtual environment while looking for the next objective.

Over the course of the experiment, participants will enter a total of 240 words (where a word is defined as 5 consecutive characters). Assuming participants type at a mean of 40 words per minute, this will constitute 6 minutes of typing. This is short enough to avoid fatigue (and short enough that typing won't seem like the sole purpose of the game), but will yield more than enough input data for later analysis. Six strings of 40 words each will be pre-calculated, and each participant will have to enter the same phrases. The strings will be randomly constructed from the 1000 most commonly used English words. Many such lists are available online. One such phrase will have to be entered to complete each objective in the game.

Real-world interactions will attempt to approximate interactions that may take place during conventional game play. However, these interactions will be worked into the game itself rather than being totally disconnected distractions from the game. For example, to complete a task, the participant will have to press a button of the correct colour. A sticker indicating the correct colour will be placed underneath a mug (prior to experiment) on the desk that the participant is at. The in-game instructor will prompt the player to reach for the cup and look underneath it in order to find out which button is the correct one to press in the game. At least one such interaction will have to be completed for each objective in the game.

The play session will take place in a controlled environment with no external distractions. This same environment will be used for every participant. Each experiment will be

[2] These baselines will be captured post-experiment, as opposed to pre-experiment, to avoid participants gleaning what some of the measures of the experiment are and any potential bias thereby introduced.

conducted using exactly the same hardware. The experimenter will remain in the venue, but will not interfere in any way unless the participant requires aid (records of any such interactions will be kept, as, in the worst case, data gathered from the session may have to be discarded).

The experiment will end when the participant has completed the game.

4.3.4 Post-Experiment

After the experiment is complete, the participant will be required to fill out several questionnaires in order to quantify aspects of their experience. Subsequently, baselines (while not wearing an HMD) for typing performance and real-world interaction will be captured. Whether or not the participant is a touch typist will be recorded. After this, the participant will be allowed to leave. Their forms will be filed, and the data gathered by the apparatus will be organized and backed up.

In order to capture baseline typing performance, a simple desktop application using exactly the same logic as that used in the game will be created. The participant will perform this typing test on the same keyboard as that used in the experiment, but it will not be gamified in any way, and a conventional computer monitor will be used. This baseline typing test will last for one minute.

Baseline real-world interaction speed will be captured by timing the participant's execution of a simple task (similar to those used in the game) while not wearing an HMD.

4.4 Measures

4.4.1 Presence

In order to measure Presence, participants will fill in the ITC-Sense of Presence Inventory (ITC-SOPI) [15] immediately after the experiment.

4.4.2 Typing Performance

Typing performance measures will be built into the apparatus itself. Various aspects of typing performance will be measured, including words per minute, accuracy, and time to first key press. In order to calculate accuracy, a modified version of the Levenshtein String Distance Statistic, as described by Soukoreff and MacKenzie [23], will be implemented.

4.4.3 Real World Interaction

After filling in the Presence questionnaire, participants will fill in the NASA Task Load Index (NASA-TLX) questionnaire [11]. This questionnaire is designed to measure work load, and will allow information about the ease of performing interactions to be gleaned. The game itself will measure time taken to complete tasks that required real-world interaction, and recordings captured during experiments will later be analysed so that the time taken to complete these interactions can be measured. This will serve as a measure of task performance. Results of this task performance will be correlated with the results from the NASA-TLX questionnaire.

5. PROJECT PLAN

5.1 Risks

Please consult Table 1.

5.2 Resources

5.2.1 Hardware

This project aims to investigate issues pertaining to VR - in particular, VR achieved through the use of an HMD. As such, a consumer grade HMD will be necessary for this research. Due to its popularity, wide user base, large volume of support, and integration with major game development platforms, the Oculus Rift is the preferred choice. The Oculus Rift CV1 (first commercial release), has recently begun shipping (28 March 2016 [7]).

In order to support the Oculus Rift, a sufficiently powerful graphics workstation is required. The minimum necessary specifications are as follows [7]:

- Video Card: NVIDIA GTX 970
- CPU: Intel i5-4590
- Memory: 8GB+ RAM
- Video Output: free HDMI 1.3 output
- USB Ports: 3x USB 3.0 ports plus 1x USB 2.0 port
- OS: Windows 10 64 bit

Finally, in order to implement the various AV techniques under investigation, a Leap Motion controller will be required. This device combines the functions of both a hand-tracker, and stereoscopic camera that can be used for video pass-through. It is small enough to be mounted to the front of an HMD[3].

5.2.2 Software

In order to create the experimental apparatus, one core piece of software is required: The Unity3D development platform. This platform has been chosen for several reasons; The researcher has previous experience with this development platform, using an alternative platform would require extra time to learn how the system works; Unity3D is widely used, there are many tutorials, and support can easily be found due to its large user base; The Unity3D development platform has built in support for both the Oculus Rift and Leap Motion controller; Finally, only the free version of Unity3D will be required.

5.2.3 People

A number of participants will be required to participate in initial pilot experiments. Thereafter, a larger number of participants will be required to participate in final experiments. The number of participants that take part in the final experiment should be sufficiently large in order for any

[3]Preliminary testing has revealed two weaknesses with the Leap Motion controller: The cameras used do not provide video pass though of sufficiently high fidelity to read the lettering on a keyboard, hampering its ability to facilitate text-entry; Because the cameras used are infra-red, all screens (including mobile phone screens) appear completely blank, rendering them unusable. For these reasons, it is necessary to use a webcam for video pass through

Table 1: Risks, Impact, and Management Strategies

Risks and Effects	Impact	Likelihood	Monitoring	Mitigation	Management
Delays in obtaining key hardware. This could result in delaying the whole project, and/or reducing the quality of the final system.	Catastrophic	7	Order tracking.	Order as early as possible.	Borrow older equipment.
Difficulties in obtaining participants. Results less likely to be significant.	High	9	Monitoring of sign-up sheets.	Power analysis to ensure I know how many participants I will need.	Reach out to friends, family, and university connections. Participation could possibly be used as extra credit for Psych undergrads.
Development delays. Could delay final experimentation and put pressure on subsequent milestones.	Medium	5	Weekly checks to ensure development is on schedule.	Detailed breakdown of sub-tasks and sub-systems required for final system created before development begins.	Sleep less.
Hardware malfunction/failure. Equipment would have to be sent overseas for repair. Potentially enormous delay to development/ experimentation.	Catastrophic	1	Weekly hardware checks.	Strict adherence to pre-established care protocol.	Fall back to borrowed hardware.
Writing delays. Could delay final hand-in.	Low	4	Weekly supervisor meetings.	Weekly supervisor meetings.	Sleep less.

potential effects to be detected with a significant degree of confidence. This number will be determined using power analysis.

Before either the pilot or final experiments are conducted, ethics clearance will have to be obtained.

5.3 Milestones

Table 2: Project Milestones

Date	Milestone
01/06/2016	Introduction and Background chapters done.
01/12/2016	Development complete.
31/12/2016	Pilot experiments conducted.
31/01/2017	Apparatus finalized.
28/02/2017	Data gathered (experiments complete)
01/04/2017	Results, Conclusion, Abstract complete.
01/06/2017	Final Draft hand-in.
01/07/2017	Final hand-in.

5.4 Timeline

Please see the proposed timeline (Appendix A), for a more detailed breakdown of the milestones above.

6. RESEARCH OUTCOMES

6.1 System

The system that will be created will be an immersive 3D game that will allow users to experience a sense of Presence. Furthermore, the game will implement an AV technique (Minimal Blending AV). In order to test typing performance while in the virtual environment, a text entry analysis system will also be built into the game which is capable of measuring all the necessary details of user input.

The key feature of the system will be the AV system that will be built into it.

The main design challenge will be the implementation of the AV system, as three separate pieces of hardware will have to work in concert. Adding enough polish to the game into which these systems are built so that users may experience a sense of Presence will be the final challenge.

6.2 Anticipated Results and Success Factors

I anticipate that the proposed AV system will result in a statistically significant improvement to typing performance, and ease of real-world interaction, compared to pure virtuality. Furthermore, I anticipate a statistically insignificant difference in levels of Presence experienced between the AV and virtuality conditions. If this is the case, it will show that minimal blending can be successfully implemented without the use of chroma-key or color segmentation, and add support to previous work which suggests that AV is a promising solution to the usability issues created by wearing an HMD.

7. REFERENCES

[1] Blizzard entertainment. http://us.blizzard.com/en-us/, 2004. Accessed: April-2016.

[2] Konami. https://www.konami.com/en/, 2015. Accessed: April-2016.

[3] Metal gear solid v: The phantom pain - no text chat in mgo3!!!! http://www.gamefaqs.com/boards/718564-metal-gear-solid-v-the-phantom-pain/72609402, 2015. Accessed: April-2016.

[4] Digicapital. http://www.digi-capital.com/news/2015/04/augmentedvirtual-reality-to-hit-150-billion-disrupting-\\mobile-by-2020/#.Vt1irrJ95p9, 2016. Accessed: March-2016.

[5] Htc vive. http://www.htcvive.com/us/, 2016. Accessed: March-2016.

[6] Metal gear solid v: The phantom pain - i miss text chat in mgo. https://steamcommunity.com/app/287700/discussions/0/458606248617248420/, 2016. Accessed: April-2016.

[7] Oculus rift. https://www.oculus.com/en-us/rift/, 2016. Accessed: March-2016.

[8] M. Billinghurst and H. Kato. Collaborative mixed reality. In *Proc. IntâĂŹl Symp. Mixed Reality*, pages 261–284, 1999.

[9] P. Budhiraja, R. Sodhi, B. Jones, K. Karsch, B. Bailey, and D. Forsyth. Where's my drink? enabling peripheral real world interactions while using hmds. *arXiv preprint arXiv:1502.04744*, 2015.

[10] U. G. Foehr. Media multitasking among american youth: Prevalence, predictors and pairings. *Henry J. Kaiser Family Foundation*, 2006.

[11] S. G. Hart and L. E. Staveland. Development of nasa-tlx (task load index): Results of empirical and theoretical research. *Advances in psychology*, 52:139–183, 1988.

[12] S. C. Herring, D. O. Kutz, J. C. Paolillo, and A. Zelenkauskaite. Fast talking, fast shooting: Text chat in an online first-person game. In *System Sciences, 2009. HICSS'09. 42nd Hawaii International Conference on*, pages 1–10. IEEE, 2009.

[13] A. Hutchison. Back to the holodeck: new life for virtual reality? In *Proceedings of the 2nd international conference on Digital interactive media in entertainment and arts*, pages 98–104. ACM, 2007.

[14] C. Jensen, S. D. Farnham, S. M. Drucker, and P. Kollock. The effect of communication modality on cooperation in online environments. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 470–477. ACM, 2000.

[15] J. Lessiter, J. Freeman, E. Keogh, and J. Davidoff. A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence*, 10(3):282–297, 2001.

[16] M. McGill, D. Boland, R. Murray-Smith, and S. Brewster. A dose of reality: Overcoming usability challenges in vr head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2143–2152. ACM, 2015.

[17] P. J. Metzger. Adding reality to the virtual. In *Virtual Reality Annual International Symposium, 1993., 1993 IEEE*, pages 7–13. IEEE, 1993.

[18] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems*, 77(12):1321–1329, 1994.

[19] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Photonics for industrial applications*, pages 282–292. International Society for Optics and Photonics, 1995.

[20] J. Peña and J. T. Hancock. An analysis of socioemotional and task communication in online multiplayer video games. *Communication research*, 33(1):92–109, 2006.

[21] T. Shibata. Head mounted display. *Displays*, 23(1):57–64, 2002.

[22] M. Slater, M. Usoh, and A. Steed. Depth of presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 3(2):130–144, 1994.

[23] R. W. Soukoreff and I. S. MacKenzie. Measuring errors in text entry tasks: an application of the levenshtein string distance statistic. In *CHI'01 extended abstracts on Human factors in computing systems*, pages 319–320. ACM, 2001.

[24] J. Steuer. Defining virtual reality: Dimensions determining telepresence. *Journal of communication*, 42(4):73–93, 1992.

[25] M. Suznjevic, O. Dobrijevic, and M. Matijasevic. Hack, slash, and chat: a study of players' behavior and communication in mmorpgs. In *Proceedings of the 8th Annual Workshop on Network and Systems Support for Games*, page 2. IEEE Press, 2009.

[26] M. Van den Bergh and L. Van Gool. Combining rgb and tof cameras for real-time 3d hand gesture interaction. In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pages 66–72. IEEE, 2011.

[27] D. Weibel and B. Wissmath. Immersion in computer games: The role of spatial presence and flow. *International Journal of Computer Games Technology*, 2011:6, 2011.

[28] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments*, 7(3):225–240, 1998.

8. APPENDIX A - PROJECT TIMELINE

Name	Begin date	End date
Proposal Complete	01/04/16	01/04/16
Introduction and Backgroun...	01/04/16	31/05/16
Design Chapter (System)	01/06/16	30/06/16
Preliminary Writing Complete	30/06/16	30/06/16
Chat Simulator	01/07/16	14/07/16
AV Implementation	15/07/16	15/09/16
Game Implementation	16/09/16	31/10/16
Apparatus Complete	31/10/16	31/10/16
Design Chapter (Experiment)	01/11/16	30/11/16
Pilot Experiments	01/12/16	30/12/16
Adjustments to Apperatus	02/01/17	31/01/17
Apparatus Finalized	01/02/17	01/02/17
Final Experiments	01/02/17	28/02/17
Data Gathered	15/03/17	15/03/17
Data Analysis	01/03/17	15/03/17
Results and Conlusion Chap...	16/03/17	28/04/17
Revisions	01/05/17	31/05/17
Final Draft Complete	01/06/17	01/06/17
Corrections/Revisions	01/06/17	30/06/17
Final Submission	03/07/17	03/07/17