VR Evaluation Techniques - A Study of Modern VR Technologies as a Tool for Experiments

Ross Anderson

School of Computing Science
Sir Alwyn Williams Building
University of Glasgow
G12 8QQ

Level 4 Project — March 12 2017

Abstract

Motivated by the recent and rapid developments in VR technology, the question of its use as a tool for conducting experiments is raised. By performing an evaluation the characteristics specific to a VE created with a 3D games engine are explored and examined. Using the data from the evaluation it was found that users were affected by the actions taken by simulated 'actors' in a manner analogous to expected results of a real-world analogue. Key limitations and characteristics were identified throughout the evaluation which indicate design constraints and considerations that must be considered when creating experimental virtual environments.

Education Use Consent

I hereby give my permission for this project to be shown to other University of Glasgow students and to be distributed in an electronic format. **Please note that you are under no obligation to sign this declaration, but doing so would help future students.**

Name: ____________________ Signature: ______________________

Contents

Chapter 1

Introduction

1.0.1 Overview

The purpose of this dissertation is to evaluate modern virtual reality systems, particularly those of low cost, in the context of providing a viable alternative to conducting experiments requiring real-world settings. This is achieved by running an evaluation using such a low-cost system It is considered with particular interest in experiments that present significant organisational challenges or face issues due to cost or safety of its participants. Key observations on the viability of current low-cost technology were identified throughout the construction and in conducting the experiment. In performing the evaluation, many features are identified which must be taken into consideration if effective experiments are to be developed in the future.

1.0.2 Advantages and possibilities of VR-based experiments

VR as an experimental tool can enable or facilitate many experimental scenarios. With the recent developments in technology, chiefly HTC's Vive and Oculus' Rift with their highly accurate motion-tracking systems, the fidelity of Virtual Environments(VE)'s has received an enormous step up from recent earlier systems. By providing a VE, one can present situations which are dangerous, expensive or abstract in a manner that allows the subject to interact with in an intuitive fashion. If an evaluation would require an impossible situation, such as scaling their movement within the world, VR rigs could allow such a scenario. Non-euclidean geometric environments, enviroments expressing concepts and other previously fantastical situations can be made available during VR. The potential for VR is not limited to enabling impossible situations, however. Impractical situations such as those requiring large groups of people might be viable within a VR evaluation. Ethical issues arise when certain scenarios are required for an evaluation. Requiring subjects to move through traffic, for example, would be difficult to achieve in reality due to the dangers and liabilities involved. VE's can provide a safe, controlled space to perform this.

VE's are also consistent. A given scenario can be constructed and run many times over a long period of time in the exact same manner without issue. Factors in real-world experiments that cannot be controlled, such as weather or onlookers' behaviour, can be controlled within a VE. The environment does not necessarily need to be in the same area at all times, either. A specified scenario can be built at relatively low cost elsewhere and the application to perform the evaluation can be transferred. This can assist greatly in organising an experiment since the VE can be reproduced on demand without requiring a dedicated space or being tied to a particular location.

There now exist 3D game engines which can be used by anyone with technical understanding to produce such environments with integrated support for these low-cost VR systems. Building-block code can enable non-programmers to script, readily available assets both facilitate in avoiding drudgework such as basic texturing as well as enabling non-artists to integrate good-quality assets. Levels can be constructed in a drag&drop fashion and rapidly produce a working environment for an evaluation. The experiment's application uses such a 3D engine in order to allow myself, primarily a programmer, to create an entire VE with actors, texturing and sound without acquiring the skillset to produce these assets on my own.

Since VR systems track the user in order to provide input to the engine, this data can be collected and automatically stored. Information such as position, timings, orientations and inertial information are all required to interact with the application. Additionally, some systems may track the eye in order to improve the image quality and so could be suitable for experiments that require eye tracking.

Considering the step forwards in VR systems and the potential for it to be used as a tool for creating experiments, this dissertation seeks to investigate such a possibility. By taking a low-cost system, an HTC Vive with a clear room and a sufficiently powerful PC, can an experiment be constructed that induces presence from social interactions? Investigating the possibility can demonstrate its efficacy. Presence for the user can be determined by their reaction to simulated stimuli. For example, if the user is arachnophobic and feels genuine fear from a simulated spider, some presence has been attained.

The experiment in this dissertation involves more subtle stimuli than triggered phobia. The effects of simulated actors performing basic interactions as substitutes for real people in a relaxed, simple, task-based exercise for the user can be hypothesised. If a group of real people approach the subject while they are performing the actions, then the expected result would be that most subjects would perform the task very differently. Similar interactions, such as having them stare at the subject and laugh at them are also included in the evaluation. These interactions are generally less stressful than in previous works, which often involve triggering phobias such as arachnophobia or acrophobia in order to demonstrate the strong response the subject feels[14]. The relative efficacy of these in terms of distracting the user from completing their tasks is evaluated.

1.0.3 Identifying limitations

These systems, while advanced, still of course have limitations. Issues with their weight and wired nature , for example, could affect the validity of results derived from a virtual experiment if they are not accounted for. Another indended outcome for this experiment is to explore the manners in which VR equipment's characteristics can potentially affect the results of an experiment. By doing so, understanding of the properties of VR can be developed. From this understanding experiments with those properties in mind can be created in the future and VR-specific experimental techniques developed.

Chapter 2

Virtual Reality

2.1 Defining VR

Virtual Reality describes a concept that has existed long before the introduction of the term itself. Of particular note is Sutherland's description of the 'ultimate display'[21]. A prophetic view of the capabilities of future technology, he describes computers as eventually possessing the technology to present a looking glass into other, potentially abstract worlds. Sutherland describes the ultimate form of this technology. One in which objects perceived by its users interact in an entirely realistic fashion including its physical presence. Sutherland also proposes more realistic extensions to 1965 technology. Sutherland puts forth a description of technologies such as positional tracking, eye tracking and force feedback as a system to construct abstract dimensions. Without describing virtual reality itself essentially the core concepts underpinning the systems to come have been made clear. To varying extents realisations of his proposed technologies have been introduced, and integrated in some form or another into our understanding of 'virtual reality'.

Virtual reality as a term has broad applicability and has been used to describe a great many technologies since. The definition can be stretched as far as applying to a 3D video - computer rendered but not necessarily generated- in which the viewer may look around, perceive limited degrees of depth but not interact in any way with the environment. Many 'VR experiences' and low-cost headsets being sold to consumers implement this. This technology sans the panoramic aspect has in fact existed since 1960 [6] in the form of the telesphere.

Despite publication in 1965, the aspects of Sutherland's vision towards a kinesthetic display saw mostly crude implementations until recently. outside of specific simulation software such as military flight simulation and training tools, attempts to incorporate virtual reality into a desirable consumer product has arguably never been successful. The most famous and understood example of this is Nintendo's Virtual Boy in 1993. The device promised customers a true 3-D, computer generated virtual environment with which they can control via the use of a gamepad. While using the device it is clear that they could not cope with the technological challenge such a promise would require using their level of technology. Sporting a monochrome display, awkward means of user input and users reporting significant discomfort during use it was quickly discontinued.

I feel that the concept of VR that Nintendo promised its users with the Virtual Boy: computer-generated environments represented in an interactive fashion with a sense of presence ('being there' in the game) most accurately describes virtual reality. Commonly amongst consumers 'VR' includes simple uninteractive 3-D video viewers however I claim that due to its lack of interactivity it can never supply a satisfactory implementation of VR. It contains part of the technology used in modern VR systems, but the display forms just one part of the whole system. The rudimentary nature of the Virtual Boy also provides context as to why VR had seen little use as a tool in its era. The level of technology required to produce usable VR would appear to be far beyond the grasp of those considering it for general use.

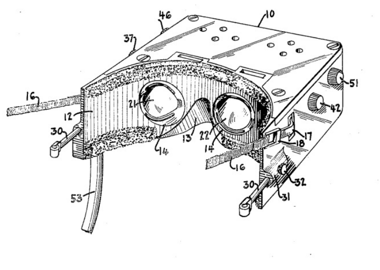


Figure 2.1: Patent drawing for the telesphere

Figure 2.2: Nintendo Virtual Boy

2.1.1 The CAVE

Despite largely falling out of commercial and public interest, VR using technology of the 90's and 2000's would still find applications. Beyond flight simulation and other specialised military apparatus, the Cave Automated Virtual Environment was commonly used to conduct VR research.[4] Originally developed in 1992 CAVE systems continued to be developed despite the lack of general public interest. A specialised room constructed from high resolution rear-projector screens, electromagnetic sensors or (eventually) infrared cameras were used for positional tracking. The tracking translated to changes in perspective of the world represented on the screens, and stereoscopic 3D vision provided via polarised 3D glasses. 3D imagery is presented in a manner not dissimilar to active 3D televisions. The projectors display a given refresh rate and the left and right eye's lens is shuttered in alternation in order to present left and right eye images. Such a solution requires a dedicated installation, considerable expense and maintenance. CAVE installations offer a few key advantages such as relatively unburdened users when contrasted with HMD's, positional tracking and generally a greater degree of interactivity with the objects. The capabilities for high-fidelity surround sound enable positional audio as well as video. The combination of these aspects can create an immersive virtual environment within which the user feels presence.

Figure 2.3: An example of a typical CAVE setup

2.1.2 Presence

The concept of 'presence' as a significant aspect of VR is argued by Sanchez-Vives & Slater(2005)[19]. They define the term and identify the key elements to attaining presence within virtual reality. They note two different approaches to defining presence. One considers it to be the sense of place - i.e. the user perceives the virtual environment more strongly than the true environment (such as the lab the experiment is performed in). Others believe presence to be defined by the user's ability to interact with the environment. The example given is that being able to walk in the environment and have that be reflected in their perception enhances the sense of presence. When using VR technologies to conduct experiments, presence is generally a highly desirable product of the experiment's design. In attaining presence, the participants are expected to respond to stimuli in a realistic manner. Meehan et. Al.(2002)[10] demonstrate the sense of place in their 'Pit Room' experiment. A virtual environment in which the user is asked to perform a basic task (placing and dropping books). The devices used in the original experiment are unknown, beyond that they use HMD's with positional tracking, but the experiment has successfully been reproduced within a CAVE environment [3].

Figure 2.4: Pit Room experiment's VE with drop shown

Eventually the user is asked to take the books into another virtual room, to which there appears to be a 20ft. drop. Presence is enhanced by requiring the participant to wear slippers and by simulating a gap in the floor using a small ledge. In a real-world version of this scenario the participant would be expected to feel vertigo and fear of injury. For many participants the effect was convincing enough that they responded in a similar fashion. Their heart rate increased and they were hesitant to step onto what they, on some level, believed to be a genuine drop. They also found the effect to be more pronounced the higher the framerate the headset displayed. Fear of falling is immediate, relatively simple to induce and evokes strong reactions. The pit room demonstrates the power of presence despite the technology's limitations.

2.1.3 Bridge experiment

With VR, there is clear evidence to suggest that spatial awareness is affected within a VE, particularly within rendered environments. Lane Phillips et. al.[15] build upon earlier experiments that found distance judgements from oneself (egocentric distance) is likely to be underestimated with a VR HMD[18]. This effect was more pronounced in entirely rendered environments than in photorealistic versions of that same environment. What this indicates is that there clear differences in the way a VE is perceived when it is rendered. The further research conducted implemented an evaluation similar to the pit involving a sudden appearance of a 'bridge' with a

20ft drop to the users' side. with changes in locale. Between the photorealistic locale and non, there was no difference in reported presence and measured stress induced. What was found was that gait changed to slower and more careful when presented with the pit's appearance in the photorealistic room. These types of factors, in which expected real-world results do not align with those generated from a simuated one, are important to consider when attempting to use VR as a tool (the revealing circumstance was architectural software in VR misrepresenting the scale of buildings) e.g. to construct an experiment. Understanding the shortcomings, being able to design around them and to what degree the technology influences the results is an important step in enabling its use as a legitimate alternative to real environments.

2.2 Current Technology

The experiment described in this dissertation uses current technology which represents a massive leap forwards in VR from systems such as CAVE or previous HMD's. 2016 saw the consumer release of two major VR headsets - Oculus's Rift and HTC's Vive. Prior to the release of these devices two major previous developer versions of the Rift were released, the DK1 and DK2. Primarily developed for video games, the Rift and Vive's HMDs share much of the same technology. As of writing both devices can be used to perform the experiment, however the Rift platform only recently received the hand controllers necessary to do so. For this reason the Rift was not considered as a development target, however it since has developed the capability to do so through the use of SteamVR.

2.2.1 Display

Both the Vive and Rift make use of a 90Hz stereoscopic display mounted behind a pair of lenses. In both cases, the HMD's make use of fresnel lenses to allow the user to focus on the close screen. Fresnel screens reduce chromatic aberrations and require less space than normal lenses for equivalent optical characteristics. The tradeoff is that areas of high contrast such as white text on black backgrounds can appear blurry. These blurry artifacts are colloquially referred to as 'God rays' - a descriptor which notes the effect's similarity to crepuscular rays. The clarity of the image reduces the further the user looks away from the center of the lens. There is an area in the center that is known as the 'sweet spot' where the image is clearest. Both the Vive and Rift have a sweetspot, and the difference in their size is marginal.

Figure 2.5: Demonstration of glare in high-contrast situations with the Vive

The Pit Room experiment was being carried out with refresh rates of at most 30Hz and a minimum 5Hz. They note the effect of presence to be more pronounced at higher framerates. The 90Hz these devices offer provide a massive improvement in this regard and should theoretically extend this. More modern CAVE systems offer 120Hz positional tracking and 45Hz refresh rates from the projectors[3]. When the Vive's rendering fails to reach 90Hz and settings to retain 90Hz such as Asynchronous Reprojection are disabled, it drops down to 45Hz. The difference between a 45Hz display and 90Hz for VR is shown to be significant, with 90Hz appearing generally smooth and 45Hz feeling uncomfortable in comparison.

The screens themselves are a pair of 1080x1200 LCD panels. The panels used combined with the lenses offer higher resolution and less blur from motion than e.g. CAVE systems which makes use of projector screens on walls, as well as a heavily reduced screen-door effect.. Between the Rift and Vive there are small differences in optical clarity in certain situations. Broadly speaking both headsets perform similarly enough that their use is interchangeable in most scenarios.

Field of View

While these VR headsets offer displays of reasonably high pixel density with low persistence and minimised 'screen-door effect' one noticeable limitation is within their presented fields of fiew. There are minor differences in the lenses and screens used for gaming VR HMD's but they all present a field of view in the region of 110 degrees total. Total field of view differs between both headsets and users however in almost all cases the field of view afforded by the HMD is significantly lower than the wearer's combined binocular and peripheral vision. There exist HMDs with larger, higher resolution screens that afford more FoV but they often have low refresh rates and high persistence[16]. This behaviour of HMDs can break established presence and affect behaviour particularly pertaining to eye movement as it encourages users to look forwards instead of naturally looking around without minimal head movement.

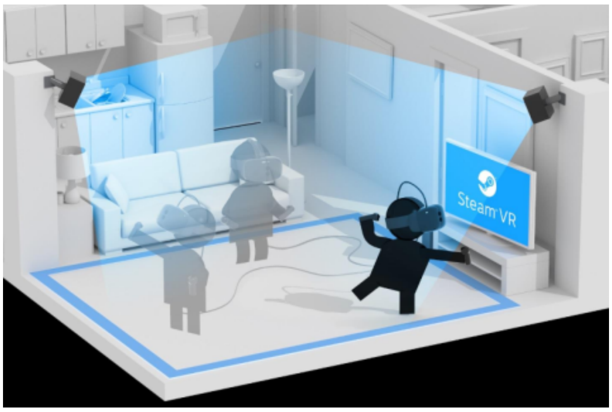
A detailed overview of the optical properties of Oculus and HTC's HMD's can be found here:[13].

2.2.2 Tracking

Optical

The optical system that the Vive uses, called Lighthouse, affords 120Hz of tracking updates. This exceeds the framerate of the displays and for context offers roughly the same rate a USB mouse will be polled on a typical Windows computer (125Hz). Tracking is achieved via devices such as the HMD or the controllers receiving signals from the lighthouse base stations. Typically configured such that one base station 'paints' the room in invisible IR light periodically horizontally and the other vertically, base stations synchronise with each other via bluetooth or link cable to interleave the flashes. The Vive has a sub-mm tracking accuracy from the lighthouse system provided the stations are not occluded by e.g. the user's body. It is a generally robust system provided the stations are installed in such a fashion that they do not shake while they oscillate. For a home installation this is most readily achieved by screwing mounts into the walls. For portable setups a camera tripod (Base stations use the same mounting system as most cameras) can create jittery tracking, and so a more stable solution such as a floor-to-ceiling monopod provides better results. Reflective surfaces such as mirrors within or around the playspace are to be avoided positioned such that the user is unlikely to occlude both stations simultaneously: typically with the stations mounted high pointing down at the center of the room.

Figure 2.6: Valve's recommended configuration for the placement of base stations



Within the experiment a portable setup was tested and successfully deployed with minimal setup time (roughly 15 minutes). Comprised of a gaming laptop, ceiling-to-floor monopods and an HTC vive an empty room could be configured to run the experiment without major tweaking. Unlike CAVE systems this affords flexibility in setting up experiments with regards to location. By installing firmware on a bluetooth dongle, one can add extra controllers to the virtual environment beyond the standard capability of 2[20]. While not used in the experiment this can be used as a quick way to model virtual objects as physical, whose realistic physical interactions can enhance presence.

The Rift operates on similar principles albeit with the devices outputting the infrared signals and cameras processing their positions to calculate distance and orientation. Named Constellation, the Rift makes use of one or more IR-sensitive cameras deployed around the player. When designing this experiment the Rift had yet to officially support any form of 'roomscale' - instead officially supporting only seated gameplay. It has since been extended with the touch controllers and 'roomscale' through optional extra cameras. Constellation does not lend itself as well to playspaces of significant size e.g. 10mx10m due to the limitations in detecting small variations in distance between points of light at a distance. The experiment in this dissertation would fit within the parameters of the Rift and Touch's tracking capabilities however it is noted that the potentially large areas lighthouse can extend to makes it capable of larger-scale virtual environments without the use of artificial locomotive controls. Theoretically it's possible to add additional physically tracked objects in the environment by attaching IR LED's to devices. This could offer a great deal of flexibility in certain experimental scenarios.

Other

In addition to the positional tracking afforded via lighthouse, the Vive includes a gyrometer and accelerometer to both assist in tracking in situations where the base stations are occluded as well as in smoothing out and augmenting data from the lighthouse system. Gyrometers and accelerometers can drift. The absolute values of position provided by the lighthouse corrects instances of drift. The gyrometer and accelerometer provide both inertial and rotational information with a high polling rate. The combination of lighthouse, gyrometer and accelerometer provide the Vive with an extremely fine level of positional data along with low latency. While the lighthouse's polling rate is comparable to the rate of modern CAVE systems, the additional data is harnessed to create a smoother experience with lower latency. These factor into enhancing the virtual experience through providing reliable responses to user input. When the user is made aware of the systems they are interacting with due to glitches presence is reduced. The experiment demonstrates some of the reliability of the technology in the lack of technical errors related to tracking, with zero instances of significant tracking loss throughout.

2.2.3 Cost

The introduction of usable consumer-level HMD's with a suite of effective, highly developed hardware has driven the costs involved with VR evaluations down significantly. Space requirements, in that a CAVE system requires a dedicated installation, have reduced down to potentially just requiring an empty room for the duration of the evaluation. With a suitably powerful PC (currently priced from £700 upwards) and an HMD with its controllers (Vive currently priced around £800) one can both develop and conduct a VR evaluation with the potential for a great sense of presence. Low-cost in the context of this dissertation refers to a complete suite of equipment priced in the region of a few thousand. There exist basic HMD's priced from anywhere between £5 (Google Cardboard) upwards. These are not considered for this evaluation because they do not include the necessary tracking equipment, rendering capabilities or fidelity. The cost of a CAVE installation could be in the regions of $1M[17]around 2008. This shows a massive reduction in the barrier to entry when considering VR for a given situation.

2.3 VR Sickness

VR sickness, also known as vection sickness, motion sickness or simulation sickness, is a persistent threat to evaluations in VR. Without due consideration it is entirely possible to construct a VE in which the majority of its users feel sick almost immediately. Research suggests that there are complicated and numerous reasons sickness may be induced through VR[9]. Vection sickness in particular draws a parallel to well-known car sickness and seasickness. While car sickness is the result of inertia failing to match perceived expectations of inertia, VR motion sicknes is often the result of the perceived expectation of inertia being met with no response whatsoever. McGill, Ng and Brewster demonstrate this by creating a scenario that deliberately mismatches motion with the perceived images, and find that nausea is reduced when some frame of reference to the motion was provided via e.g. peripheral vision. Lien et. al.[8] find repeated exposure to motion sickness-inducing scenarios can develop a resistance to its effects, but not in all cases. The majority of the general population(75%) will experience normal levels of motion sickness. The analogues to regular motion sickness do not cover all the reasons for discomfort during a VR experiment. The results of the experiment identified potential additional issues that can cause discomfort. Generally recommended practices from to reducing the likelihood of VR sickness in as many participants as possible include:

- Maintain a stable framerate of at least 90Hz. Particularly avoid the possibility of a sudden drop.
- Avoid artificial locomotion entirely

- Avoid situations in which the player experiences acceleration
- Limit the length of VR sessions
- Increase field of view and visual fidelity of the headset

More specific practices can be found within Oculus' Best Practices guide [12]. Sudden framerate drops are difficult to avoid, particularly with quickly designed systems, but are essential in conducting an experiment where VR sickness would interfere with the results. Consideration should be given to experiment length. When designing VE's (Virtual Environments) for the Vive, the simplest approach to minimising the risk in addition to maintaining framerates is to make heavy use of Roomscale. By designing situations in which the participant is not expected to perform abstract tasks such as pointing and clicking the touchpad to move Artificial locomotion, methods of moving through the VE in a manner similar to traditional first-person games, is a hot topic particularly amongst VR game developers. The recommendation to allow users to instantaneously teleport to various locations has significant gameplay implications and as such there are countless attempts to include artificial locomotion in games whilst also minimising the potential for sickness. Oculus include a 'comfort rating' assigned to the games on their storefront which indicate the likelihood of VR sickness during play. Games introducing artificial locomotion seems to always induce VR sickness for some portion of users which indicates the dangers of deploying it within an experiment. Of games making use of artificial locomotion, the first person shooter Onward's implementation stands out in terms of the low reported levels of sickness experienced. On the other hand, the game Windlands, in which the user swings via ropes and artificially climbs buildings, stands out as an example of a game which is likely to induce VR sickness. While many of the mechanics of VR sickness share circumstance with general motion sickness, VR sickness specifically is not fully understood which is why publishers recommend such specific practices to avoid VR sickness, and why many independent game developers struggle to provide locomotion within a VE while also attempting to follow best practices.

Chapter 3

Design of Experiment

The previous chapter illustrates the rapid and recent developments in VR technology. These developments introduce hardware with a different set of physical and practical limitations. To successfully use the new hardware in evaluations the simplest way to refine the process is to perform evaluations themselves and reflect on the results. From this an evaluation of the suitability of modern hardware as an analogue to real devices can be produced. This chapter describes the rationale for the design of the experiment conducted in terms of what the various aspects to be included aim to achieve.

3.1 General Outline

In order to develop an evaluation, an experiment outline was drawn which was then implemented as a 'game' in the Unity engine. The experiment's outline involves requiring the user to perform the same set of simple tasks in a variety of controlled scenarios. I decided to take advantage of the Vive's advanced tracking and haptic capabilities and have the user perform various 'gestures' in sequence. During the experiment the subject is requested to place the appropriate controller in the appropriate location. Gestures are generated via the transitions of states between locations.

While the subject performs these actions, various social interactions occur with varying numbers of simulated actors. These interactions should occur at random (with the cause left open to interpretation by the subject) and may incur a response within the subject that affects their performance.

The process of performing simple tasks and introducing stimuli externally was inspired primarily by the pit room experiment[10]. This structure provides distraction to the participant in terms of what they should be doing, which provides enough activity to prevent boredom while not being so distracting that it would be impossible to focus on other tasks. The tasks chosen should be simple to perform and not require significant physical exertion. The gestures chosen were designed such that they could be comfortably performed and allow the participant to simultaneously observe the surroundings.

3.2 Task

The task chosen was an abstract form of movement exercise. The inspiration from this came from looking at the capabilities of the controllers and the relative ease with which the task could be communicated to subjects of any background. All the subject is required to do is hold the correct coloured hand in the correct coloured ball.

When held, it communicates via the combination of a tick, a pulse and a small haptic pulse every second, for 3 seconds, on the controller that is being held in the correct position. When the balls are held long enough, they make a sound and clearly move to their new position.

As most subjects will be unfamiliar with the VR controller used and many rarely playing games, the use of buttons was avoided. This reduces tutorial time, setup time, and the likelihood the subject will have their experience interrupted if they forget the particular button required to proceed. In addition to this models of hands represent the controllers in-game. These have no visual indicators of the positions of the buttons on the controller so the subject is extremely unlikely to know which button to press without prior experience to using the Vive's controllers.

Figure 3.1: User's representation of their hands

Figure 3.2: Users hold their hands in the balls to progress

3.3 Response

Meehan et. al. suggest further research along the lines of their own sugesting the evaluation of less extreme forms of stimuli than the pit room. To some extent I feel the technology present in these earlier experiments did not have the fidelity required to evoke response beyond extreme stimulus and as such attaining strong presence with the technology difficult to achieve. The advancements in VR have opened an opportunity to re-evaluate presence with a VE. Fear of falling is well known and intense response. I decided to look at an evaluation that would involve social pressure applied via simulated 'actors' and consider their presence and actions' effect on the performance of a basic task.

3.3.1 Eyes and observations

Early on, the feeling of being watched was identified as a feasible element within the experiment. The mere presence of eyes have a well-understood effect on behaviour[5]. Ernest Jones et. al. demonstrate that an abstract presentation of eyes (on a poster) increased the rate of socially altruistic behaviour (clearing one's table after lunch at a canteen). Burnam & Hare's research reinforces this concept through a public goods game in which 'Kismet' - an anthropomorphic character with eyes, observes some players' contributions to public goods contrasted with keeping a portion of the money to themselves.[2] For these reasons I decided to have the actors stare at the user at points throughout the experiment. Kismet demonstrates recognisably non-human 'observers' can influence behaviour, while the effects of images of eyes themselves can do the same. The question of simulated participants in the experiment and their influence on the participant in this dissertation's experiment is raised, and the feedback of the subjects used to answer it.

3.3.2 Crowd

Experiments involving crowd anxiety have been conducted within the context of VR. Taffou et. al.(2015) [22] performed an experiment using CAVE-like hardware investigating the effects of crowds on agoraphobic individuals. Of particular note is their reported VR sickness issues, with 1 experiencing it to such a degree they could not complete the questionnaire, and with others reporting some degree of sickness. They found discomfort increased with the crowd close to the subject. Introducing the elements of crowds and proximity was included in our experiment for this reason. Their analysis of their experiment suggests the technical constraints of the hardware posed significant problems throughout the experiment, many of which I believe can be overcome with the Vive.

Effects behind

A simple situation, where if in the real world would provoke a predictable response, is that of cues from surprising locations. In the experiment, the effects of placing actors behind the player were considered and their behaviour to their presence monitored. When the subject is made aware of the presence of actors behind them, they may feel uncomfortable performing actions that require them to look away. This could result in a variety of interesting behaviours.

3.4 Hands

In order to communicate that certain hands are to be placed in certain balls, the hands were coloured orange and blue. These hands took the form of basic shapes with bright colours and included thumbs that indicated left or right. Since the software tended to swap the hands' positioning when programs were run, the subject was asked to swap controllers around if the hands looked incorrect. Since the Vive by default supports only 2 controllers and a head position, representations of hands was considered the best choice in allowing the user to feel like actively within the VE. As outlined by Sanchez-Vives and Slater [19], virtual body representation can enhance presence but also be detrimental if there is too much that doesn't align with one's body, it becomes disorienting. The hands provide some measure of familiarity and are reasonably close to the true position of their hands from their own perspective. If entire arms or bodies were to be included in the player model it would be unlikely to work without additional tracked points of reference (such as arm/leg pads, tracked belt etc.) for the model's rigging. With just the positional information from the two controllers and the head the best approach was deemed to be to just include hands without additional virtual representations as in 3.1.

3.4.1 Audio

Experimental setups for VE's often involve sound as an element. Spatial elements such as distance or sound can help construct a convincing environment. In the interest of evaluation spatial audio within Unity and the Vive, I decided to include laughter as a social interaction. The laughter used should not be particularly taunting, however the user may perceive it to be in response to their performance and be affected in some way by this. The audio source within the engine was set as an attribute of the actor itself - meaning that it should be positionally located where they stand. In addition, the hardware used during the experiment should be sufficient to provide accurate positional audio. When the actors laugh or cheer behind the player, the expected response would be for the player to look behind them to see what caused the noise. It also should provide a cue for the player to realise that there are actors behind them.

3.4.2 Environment

Figure 3.3: Typical initial scene setup

To reduce the possibility of distraction but also retain familiarity with the environment, a basic room with wooden flooring, white wallpaper, simple lighting and a window looking out to generic buildings was chosen as the environment to perform evaluations with. The design intent was to include just enough detail that it could be understood as a simple model of a yoga room or flat to provide some context for the gestures to be performed. The window was placed facing out to make the room feel less confined and natural. A virtual yoga mat is present in the center of the room. This serves to communicate the center of the room as the location for the activities and to provide the subject with a quick reference for where to return to if they choose to step away for any reason. The yoga mat and gesture balls were positioned in such a way that standing on the mat and performing the gesture encouraged a consistent orientation with regards to the walls.

Chapter 4

Development

This chapter covers the decisions and process behind developing the Virtual Environment the experiment uses. It justifies early decisions, outlines the process and reflects on the experience.

4.1 Platform

When developing software involving 3D graphics and environments, there exist a wide range of options. At its most bare, OpenGL provides low-level access to 3D hardware. Typically involving C++ development, OpenGL gives the most control over the hardware. At the highest level of abstraction, engines such as Unity and Unreal Engine 4 provide significant structure and easy-to-use API's which assist greatly in producing a working product. There are numerous game engines e.g. CryEngine, Source that lie between these extremes. Unity and UE4 are of particular interest to those looking to develop VR games as they offer immediate API access to Valve's OpenVR SDK. An example of an issue one may face when developing for VR is in how it reacts to objects being 'thrown' by the player. OpenVR and the VR Toolkit immediately provide sensible default kinematic behaviour in these events (objects inherit velocity both rotational and linear as the user lets go and interact with others according to their weight). Despite Source Engine being a Valve product, even they chose Unity for their 'The Lab' game - a series of Vive demos available on release.

Both Unity and UE4 allow the developer to code their game's scripting in a high level fashion. Unity offers both C# and Javascript while UE4 offer C++ and Blueprints, which are structured, visual building blocks of code which can be constructed in C++ that interact in an intuitive manner. Unity offers an extension, PlayMaker, which enables similar block-based visual scripting.

One of the biggest advantages of choosing a high level engine is in the support for software asset packages. The success of these platforms amongst independent developers along with the high level scripting support is in the ability to include and purchase, all within the engine itself, pre-built assets for rapid development. This approach can enable independent developers to procure basic assets and scripts to accelerate development and fill roles. Since the evaluation was to take the form of a 3D environment in which the user is expected to perform game-like tasks, much of the challenges an independent game developer faces applies also.

Ultimately, Unity was decided as the platform of choice. Platforms that support VR natively only were considered, which at the time of decision was UE4 and Unity. Unity's VR support seemed more developed and community supported with the developers of the headset itself also providing support. The VR Toolkit offers guidance and example scripts and levels to understand the interaction with the Vive's hardware. As I do not have the experience to implement these VR functions manually and efficiently, I picked Unity and C# as the platform to develop the experiment.

4.2 Process

During the development of the software to run the experiment, sets of requirements ranging from 'must-haves' to 'would-haves' (MoSCoW requirements) were categorised and prioritised. Due to my relative inexperience with developing games I made rough approximations of the effort required in implementing the tasks involved in meeting these requirements and based development priority on these approximations. This, for the most part, worked well. There were instances of tasks I believed to be of great difficulty (such as implementing the actors with animations) that were completed quickly (mostly due to the Morph3D system). On the other hand, the length of time to implement the gesture system was heavily underestimated.

4.2.1 Programming

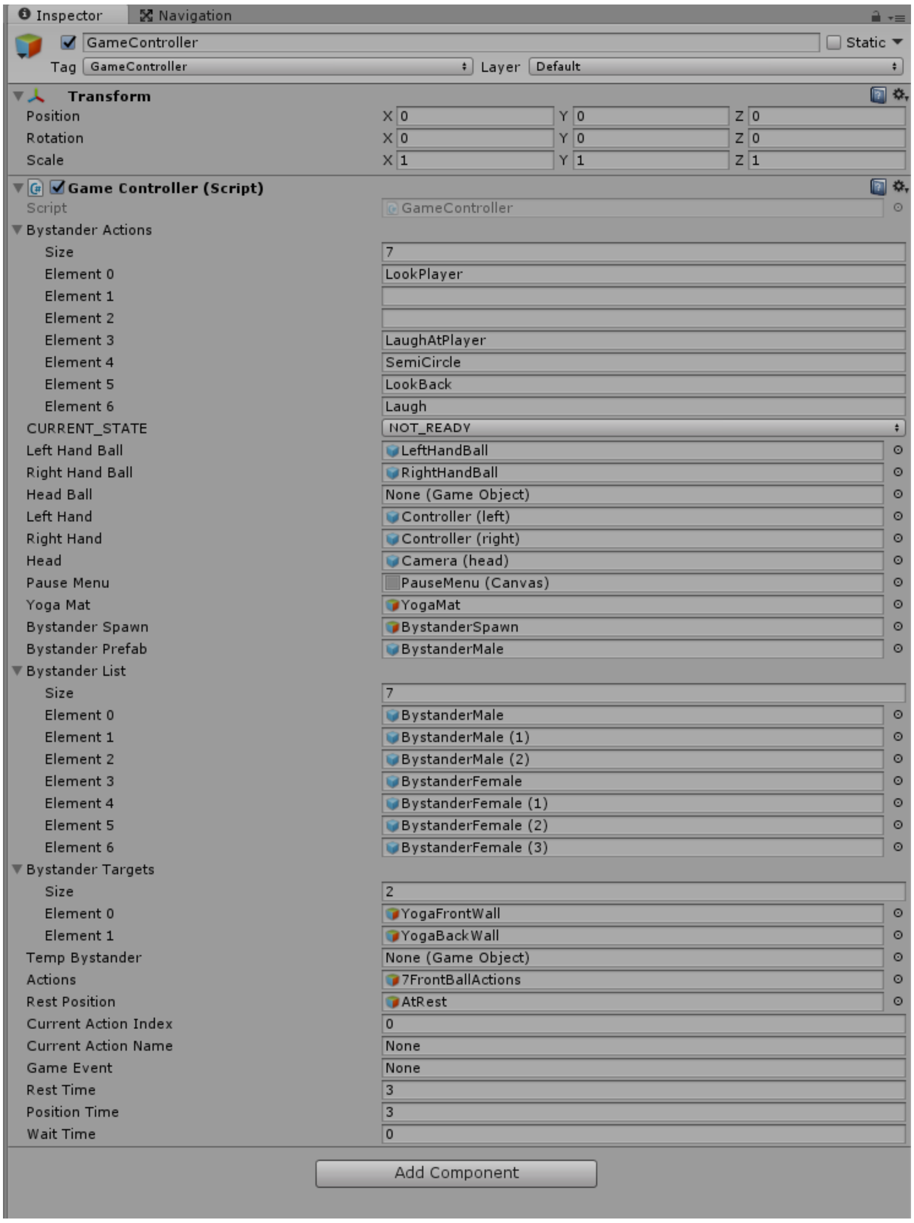
Between JS and C#, C# appeared to be the language of choice for most tutorials and experienced Unity developers providing online assistance through the Unity forum. Combined with my familiarity with Java and C, C# was chosen as the language in which to write the scripts that drive the behaviour of the game.

Despite C# being primarily an object-oriented language, Unity makes use of C# scripts in a component-based architecture. It technically can support fully object-oriented design, however the effort required to implement it within scenes increases significantly compared to building components separately.

For example, when a member of a class is declared in a script as public, OOP would require instantiation of the member through the use of an explicit constructor method. This does not necessarily apply to Unity. Fields made public can be seen in the menu options for that particular object, and a component matching its type can be dragged in. A good example of this within the program's code is in the various 'levels' (which represent the various configurations of actor size and placement). The same game controller script is used in all but the tutorial, with various changes made in-scene by dragging the relevant actors, actions, gestures etc. into the controller script. The power of this approach to an evaluation's design hierarchy is in the ability to generate specific controlled instances of a base level by simply dragging and dropping the relevant components into the scene.

Actors can be placed manually with the scene editor and their reference dragged and dropped into the controller script. This approach is much more intuitive when attempting to tweak the parameters of a given instance than e.g. the OOP approach of instantiating the various objects with a constructor script or subclassing the game controller with manually configured behaviour. The rough layout of the level could be seen directly in the scene viewer which allows for a more intuitive process of design than manually scripted actor placement at runtime.

Figure 4.1: Game controller with exposed parameters to drop objects into



Recorded Data

Since Unity's control structure operates with the game loop design pattern, in which calculations, rendering and checks are performed in a continuous loop, data from the engine was acquired in the form of snapshots of game state via the game loop. Logging was done on a by-instance basis with unique filenames all representing a particular level at a particular time. Data was collected in a logging function and used to generate a csv file. While not all data was used in the analysis, in-game positional values of the HMD and controllers, a timestamp, the state the player was in, the action they were currently performing and instances where actor interactions occur were captured as a row and saved to hard disk at the end of the level. The loop in which the logging occured ran at roughly 90Hz (framerate drops can affect this) - which provides as many snapshots per second. The csv data the application produced could then be exported to data processing applications for interpretation.

4.2.2 Assets

The majority of assets used within the experiment: textures, sounds, models, animations, and shaders were obtained online. All assets are either available under a free license or derived by myself (scripts, specific characters, level structure etc.). Sourcing assets from the asset store allowed me to quickly generate a suitable virtual environment without learning a wide array of skills such as 3D modelling , graphic design or animation. Unity's asset store contains many generic assets e.g. trees, basic buildings, wood and wallpaper textures. integrating these with the project was quick, simple and produced results far better than what I could achieve on my own. After downloading or purchasing the asset it is made available as a component. These can be dragged directly into the game world and immediately used. I feel this type of game engine, one in which 'boilerplate' assets such as trees can be obtained and integrated with minimal effort, can be an extremely valuable tool in constructing VR experiments, allowing the researcher with missing skillsets to avoid low-quality solutions (such as basic cubes representing a tree) or significant time investment and focus on developing an effective environment.

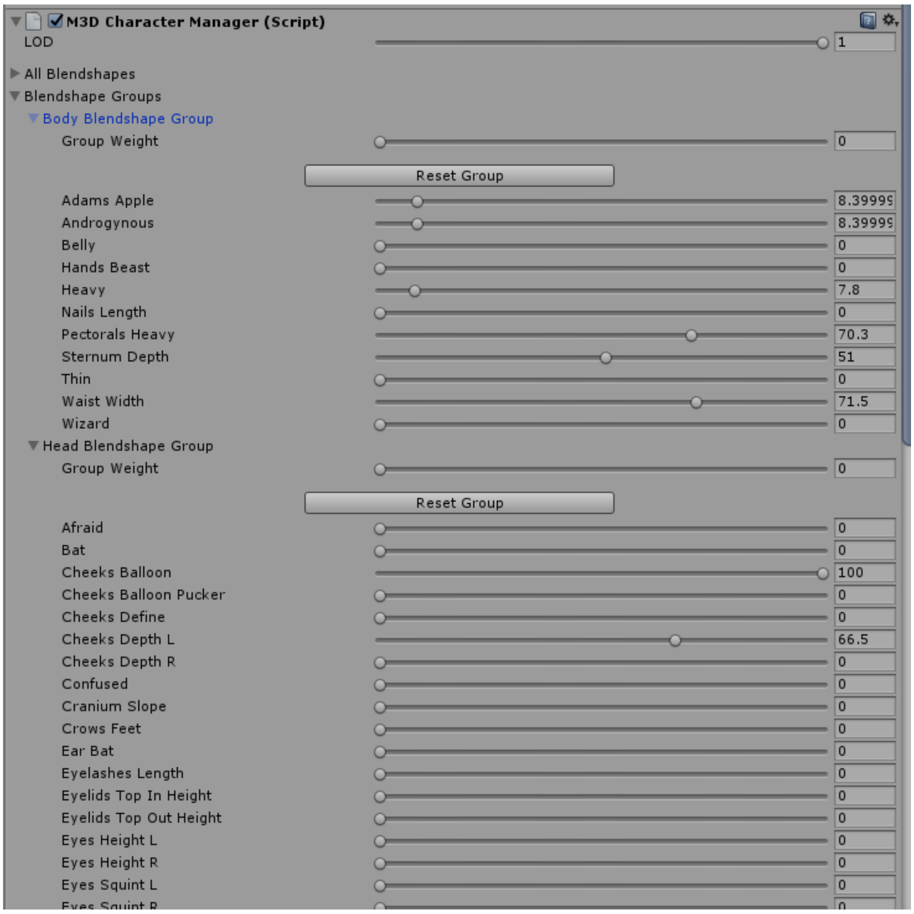
Sound

Sound clips were sourced from freesound.org. Freesound is a repository of free-licensed audio samples which provided me with the necessary beeps, laughter sounds etc. to incorporate the audio feedback necessary. Within the levels, the sounds were 'played' virtually from the virtual space that corresponded to its source e.g. beeps played 'from' the ball or laughter 'from' the actor.

Morph3D

Actors present within the experiment were generated using an asset package, Morph3D. At reasonable cost (roughly $40 value total in this experiment) an engine to produce characters with a wide range of animation and default behaviours is provided. Using a slider-based system with swappable clothes a variety of characters with reasonable fidelity was produced. The provided animations from both the Unity engine and Morph3D itself allowed me to include non-static characters within the evaluation, which is significant in terms of attaining presence as the actors can interact with the player in a realistic fashion. For researchers looking to add human characters into their evaluations I would recommend the Morph3D system or similar prepackaged systems as a shortcut compared to manual creation.

Figure 4.2: Example of sliders which generate the actors in Morph3D



4.3 Testing

While developing the software I used my own home VR setup to test the software. Consisting of a desktop PC and an HTC Vive I would run repeatedly through the simulation in order to identify bugs. Of particular note is frame dropping. The experiment requires a stable, high framerate in order to reduce the number of possible reasons the subjects might experience VR sickness. During testing it was found that very large crowds (10+) were likely to cause frames to be skipped as they moved around in the scene. By reducing the polygon count of the actors and the number of actors in the 'crowd' instance of the scenario it was possible to maintain an entirely stable 90Hz consistently across iterations. Larger crowds were desirable, and it may be that some optimisations or a more powerful PC could enable the original design to run consistently.

4.3.1 Portable Setup

While ultimately unused for the experiment since I found it easier logistically to conduct the evaluations at home than at the university, the target platform for the experiment was a portable setup. Containing 2 floor-to-ceiling monopods, a Vive, its base stations and controllers and a powerful laptop, the entire setup was shown to be capable of being packed and carried by one person and deployed and calibrated within around half an hour. It was shown to work in an empty room without issues with tracking. Since there were concerns with performance in moving from one PC to the laptop it was tested on multiple occasions throughout development to ensure the performance would not drop below 90Hz if the portable setup is used.

Chapter 5

The Experiment

5.1 Experimental Procedures

5.1.1 Sample

10 subjects were evaluated with the same procedure. Of varying age between 20 and 65, 7 male and 3 female subjects. Subjects were asked about their experience with video games, VR, and exercise routines such as Yoga, Tai Chi or Karate prior to the experiment. Of the 10, 4 reported no experience with VR and minimal experience with video games. Of the 6 who have experienced VR, all of them had experienced it through the Vive and only 1 had sigificant time with it (over 2 hours).

5.1.2 Experimental configuration

Figure 5.1: Vive HMD and controllers used in the experiment

Figure 5.2: Environment with user performing a gesture

All evaluations were performed with the home setup as pictured. This provided a roughly 2.5mx2m designated space to move in. With this area and a generally stationary evaluation, the subjects had enough room to perform the experiment without feeling constrained by the walls. The chaperone boundaries which appear as a grid of lines when a device is detected within 30cm of them did not appear for any significant length of time. My personal vive has one modification from standard which is the strap for the HMD. As shown, it uses a rigid frame with a custom mounting bracket instead of the standard flexible elastic. This proved to be beneficial for users placing it on their head as the strap can be adjusted by themselves or easily assisted with. The default strap does not readily adapt to different head sizes and face shapes, and the non-rigid structure can cause a poorly adjusted strap to cause sagging and reduce optical clarity.

Headphones used were Sennheiser HD598's which are comfortable, produce a wide soundstage for positional audio and are of an open design, allowing the wearer to speak freely and listen to instruction or ask questions without the assistance of a microphone.

Subjects wore a Microsoft Band 2 throughout the experiment with the intention of tracking GSR and Heart Rate. There were significant issues with the data produced by the device stemming from issues such as slipping up and down the wrist throwing off the readings and with the GSR's potential for useful information overwhelmed by noise in the results caused by the movements required to perform the gestures. The data from the Microsoft Band was found to be unusable and discounted.

The PC used ran Windows 10 with an intel i5 4690k processor and an nVidia GTX 970. With this PC a stable 90Hz was achieved throughout the experiment.

5.1.3 Experimental Procedure

Prior to the experiment, IPD was measured with a ruler, then the value set on the HMD. This value, a measurement of the horizontal centerpoints of the lenses, is significant for comfort during the experiment as an incorrect value can cause eye strain when used for any significant length of time.

Each subject was put through a short training course as both an equipment check and to ensure they understand how to perform the gestures to proceed. The tutorial contained relatively large written instructions on a floating scroll. If the user could not comfortably read the tutorial text, they were asked to remove the headset in order to adjust settings or take on/off glasses. As well as instructing the user on how to perform gestures, the test provided a consistent minimum level of optical clarity when adjusting the equipment for the particular user. The tutorial level has no actors present.

With the tutorial completed the subject was asked to perform each set of tasks as they were loaded into the environment. The program was structured so that each scenario with significant change was its own level. The levels were run through sequentially, with the ordering of the sequence randomised on a per-user basis by shuffling cards. The subject was not made aware of the order or the nature of each level beyond that it will require similar gestures to be performed to complete and that they would be running through 5 iterations.

Each level was configured at run-time with a discrete set of gestures. The order of these gestures were randomised. The experiment was designed such that all subjects would perform the same set of actions and the same set of levels with an equal number of actor interactions but never in the same order. To allow consistent transitions, a gesture consists of a rest position, a start position and an end position which then returns to the rest position. Without a rest position there ran the risk of creating very short or long gestures due to its random ordering.

In levels with actor interactions enabled, interactions were randomly assigned to the mid-point of a specific gesture. 1 pre-coded interaction was assigned to 1 gesture at random, with the remaining gestures given no interaction. Interactions included:

- All actors turn to face the position of the player, continue to 'stare' for a short time and then return to their previous position.
- All actors laugh. Sound played is consistent with the number of actors
- Same as above except they first look at the player then laugh
- Actors face the wall away from the player
- Actors walk towards the player, stop, stare at the player for a few seconds then walk back

The levels were configured as follows:

- 2 Side: 2 actors, 1 male 1 female, placed to the right of the subject, looking out the window. They stand with an idle animation and do not interact in any way for the duration of the level. 2 Side is the only iteration other than the tutorial in which there are no actor interactions.
- 2 front: The same actors as above were placed roughly 4 meters in front
- 7 front: Same as above with 7 actors.
- 2 back: Same as 2 front but with actors placed roughly 5 meters behind the user

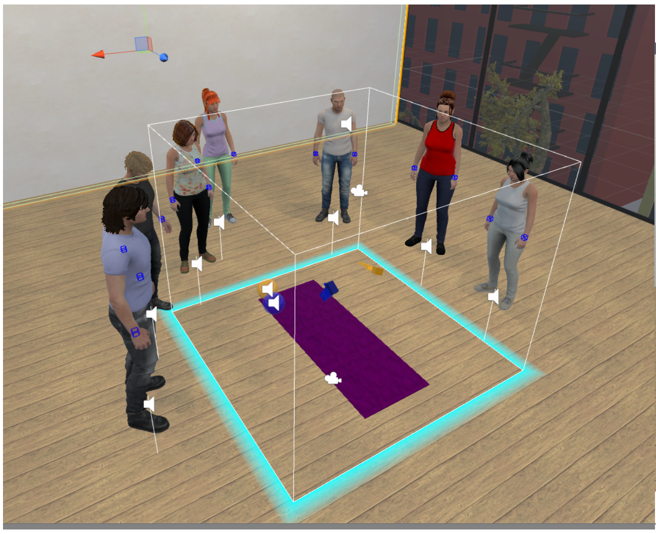


Figure 5.3: Interaction where the actors walk up to the player

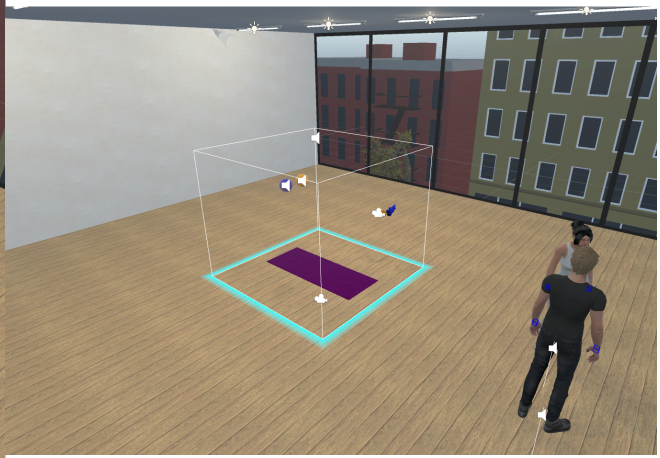


Figure 5.4: Room setup with 2 interactive actors behind the player

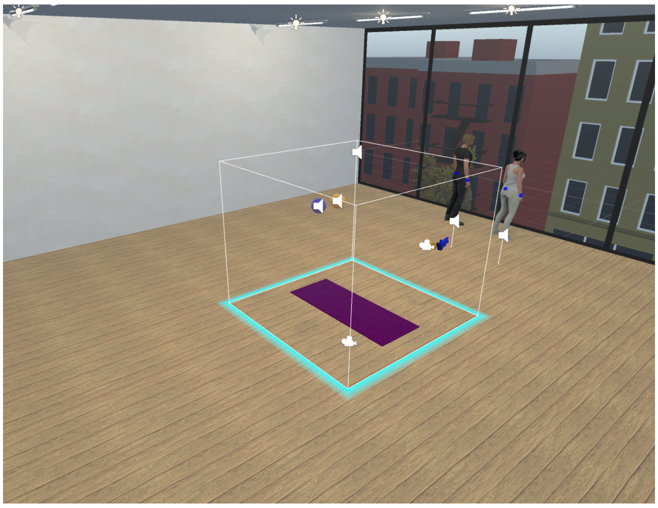


Figure 5.5: Room setup with 2 actors to the side and uninteractive

While the subject performed the experiment, comments, unusual behaviour, and issues with the experiment such as in the event of a bug or crash were noted. Observations regarding their response to the interactions or issues performing the gestures were also noted.

After the subject ran through all 5 iterations, they were asked to remove the headset and conduct a short interview. This interview was conducted verbally with the responses recorded and sorted. All subjects were asked the same set of questions, as follows:

- Did you have any issues performing the gestures, particularly with respect to the physical placement of the balls?
- Do you feel ill or carsick either while performing the experiment or now?
- In each instance, when you noticed people, how many men or women total did you see?
- Did you feel the presence of the actors affected your performance?

- Were any of the actions themselves particularly distracting?
- Did you find the actors to be more distracting when the actors were placed in front or behind you?
- For subjects who did not appear to notice actors behind: ask if they were aware of them, and general comments on those situations
- When laughter was heard and the actors were not in front of you, were you able to tell that the sound was coming from behind you?
- General comments on the experiment itself

Chapter 6

Results

6.1 Quantitative

While observing the participants, notes were taken corresponding to their behaviour. particular attention was paid to whether they appeared to take notice of actors, and whether they looked back at them after noticing. Recorded by the program are the timestamps for the particular activities. Various logging data was exported to csv by the unity application. Removing erroneous data from the results (when one user had an issue understanding the presentation or execution of a gesture the timings were omitted). There additionally was 1 crash in the first iteration (after 3 gestures) of one participant's recordings. The iteration was restarted, old data removed, recorded and included.

6.1.1 Completion times for scenarios

Taking the average value of each scenario's completion time from the moment the first action is started to the moment the last action was completed:

Figure 6.1: Full timings for given iterations, notes and averages

	2side	2front	2back	7front	7back	Total	Iteration order	Notes on recordings
Ra	85	90	80	87	87	429	2s2f2b7f7b	
Ry	84	87	88	97	76	432	7f2b2s7b2f	
Ia	85	109	100	100	112	506	2s7b2f7f2b	
Ha	84	93	82	92	82	433	7b2f2b7f2s	
Ma	86	88	84	89	80	427	2s2f7f2b7b	
Na	81	84	78	86	77	406	2s2f7f2b7b	
St	74	77	74	78	76	379	7b2f2b7f2s	Program crashed shortly into first iteration, discarded and restarted.
An	129	131	122	128	126	636	2b2s2f7f7b	Application bug, 5 additional gestures (no bound actor interaction) per iteration
Ja	83	90	95	N/A	85	N/A	7f2b7b2s2f	Difficulty understanding gestures in first iteration (7F), outlier
Wi	78	79	73	89	76	395	2s7b7f2f2b	
Average:	86.9	92.8	87.6	94	87.7			

Table 6.1: Mean times across all 10 participants for scenarios

Iteration Name	Mean Time(s)
2 front	92.8
7 front	94
2 back	87.6
7 back	87.7
2 side	86.9

6.1.2 Relative per-task performance with actor interactions

An interesting interpretation of the CSV data is to look at the average time taken to complete tasks with a given type of interaction in a given type of scenario (those with actors in front and those with actors behind) and compare it to the average of the same task performed in the same type of scenario but with no interaction. What this can tell us is the impact of the actor's interactions on the player relative to one another with respect to task time: Tasks are gestures with specific names that either have a generic identifier (e.g. ActionA) or one that roughly describes its action (e.g. Lunge)

Figure 6.2: Average actions and timings for rooms with actors in front

	Laugh		LaughAtPlayer		LookPlayer		SemiCircle		None	
Player Action	Count	Avg	Count	Avg	Count	Avg	Count	Avg	Count	Avg
ActionA	2	12.83	3	15.68			1	13.46	4	12.07
ActionB	2	11.66	3	12.24	2	14.77	1	13.20	2	14.10
ActionC	2	11.36			1	11.54	1	11.96	4	11.67
ActionD			3	12.02	2	11.03	3	11.08	21	11.36
ActionE							1	17.76	1	13.53
ActionF	3	12.51	1	11.86			2	15.47	12	13.03
ActionG									10	12.10
ActionH	1	13.40	1	11.54	5	18.79	1	14.83	2	12.38
Cheerleader	3	11.65	3	13.82			1	21.82	1	12.48
Cross	1	12.91			2	12.16	1	11.87	14	12.17
Hadoken	2	14.64			2	11.18	3	13.76	7	12.33
Hidden Dragon	1	11.12	1	10.61	1	12.49	3	15.88	12	11.23
Lefty Lunge	1	10.17	2	11.72	2	10.94	1	19.69	3	12.78
Paint The Fence									10	11.05
Reverse Starman	2	14.87	3	13.69	3	14.61	1	11.44	2	14.67
Starman									1	11.79
Grand Total	20	12.54	20	12.99	20	14.10	20	14.46	106	11.99

Figure 6.3: Average actions and timings for rooms with actors behind

	Laugh		LaughAtPlayer		LookPlayer		SemiCircle		None	
Player Action	Count	Avg	Count	Avg	Count	Avg	Count	Avg	Count	Avg
ActionC			1	16.10	3	11.52	2	10.80	3	11.36
ActionD			1	10.94			4	11.17	14	10.95
ActionE	5	11.92	2	14.26	1	10.21	5	11.92	5	11.94
ActionF									10	12.75
ActionG	2	12.67	2	11.55	1	11.32	2	13.00	18	12.62
ActionH									1	11.20
Cheerleader	3	12.13	2	11.16	1	10.96	1	23.31	4	11.54
Cross	6	13.19	5	12.63	4	13.25	1	11.44	11	11.96
Hadoken			1	12.31						
Hidden Dragon	3	15.26	2	10.37	2	10.98			12	11.09
Lefty Lunge	1	9.92	2	10.96	4	11.22	1	10.40	9	12.00
Lunge					2	12.03	2	11.14	4	11.64
Paint The Fence			2	10.62	2	12.33	2	11.04	14	11.13
Reverse Starman									1	14.58
Starman									1	11.48
Grand Total	20	12.81	20	12.02	20	11.78	20	12.07	107	11.78

6.1.3 Turning Around

A surprising result of the experiment was in the number of users who never turned around to see what was behind them. In the apparent absence of actors, despite hearing the laughter, many ignored the interactions and continued the gestures. They occasionally turned round at the end of an iteration (where there is a 5s wait before the level exits) but during the gestures the environment was completely ignored. This even applied to the 2 side iteration, in which the actors were placed directly to the right of the typical orientation during the gestures. Of the 10 subjects, 5 of them looked back. Of the 4 with no prior VR experience, 3 did not look back. of the 6 with experience, 2 didn't turn back.

6.1.4 Age

Age does not seem to correspond to completion time within this evaluation. The sample included 2 65+, 2 50+, 4 20+ and 2 30+ individuals. Placement of gesture balls were such that all users could easily reach them without requiring them to bend over or stretch. With one exception (a 'fencing lunge' move) users could perform the gestures without requiring to take a step.

6.2 Qualitative

Qualitative data was obtained through observation of the subjects' behaviour and comments. In addition, the questionnaire after the experiment allowed for participants to express their thoughts on the experience and describe aspects that were not apparent externally.

6.2.1 General behaviour

- 2 participants appeared to completely ignore the environment beyond the gesture balls. They quickly completed each gesture as fast as possible, never turning around.
- One participant, upon becoming aware of the 7 actors behind him, walked around to the other side of the balls. He performed gestures by crossing his arms (to match left and right) and watched the actors while doing so.
- Some participants appeared reluctant to move from their starting position. 3 were noted as having a strong preference for stationary feet and were reluctant to take the step forwards to lunge.
- 5 users were clearly distracted when the actors they were aware of moved towards them. They showed this through behaviour such as stepping back, turning to look or commenting on the actors' behaviour. This often caused their hands to drift out of the gesture they were holding, and require them to restart the counter.

6.3 Interview

Each subject was interviewed for no more than 5 minutes using the questionnaire. Informally conducted, the answers for the questions helped clarify the subject's engagement with the task and provided evaluation on the gestures themselves. The general sentiment of the responses will be collated.

Did you have any issues performing the gestures, particularly with respect to the physical placement of the balls?

6 users did not consider any actions difficult. The other 4 found it difficult to maintain the hands in the position while they were far apart. They could not see both simultaneously and drifted out of the balls.

Do you feel ill or carsick either while performing the experiment or now?

No users reported VR sickness throughout the experiment.

Did you feel the presence of the actors affected your performance?

2 users felt they had little effect on their performance. The remaining 8 agreed that it affected their performance to some extent. One user claimed their presence motivated them to perform better, others claimed it was distracting.

Were any of the actions themselves particularly distracting?

3 remarked on the laughter. 3 users remarked on the 'stare'. 4 users commented on the walking towards interaction and 2 felt they were unaffected. Additionally, 1 was distracted by the 2 side iteration's actors.

In each instance, when you noticed people, how many men or women total did you see?

Subjects generally were unable to accurately or confidently state the size of the crowd. Answers regarding the large crowd ranged from 5 to 8, however in all cases it appeared to be a rough guess. All but 1 stated they saw 2 people in some iterations. The remaining person paid no attention to the actors and could not give any real estimate in any iteration.

Did you find the actors to be more distracting when the actors were placed in front or behind you?

Note: If the subject did not notice actors behind them ask for speculation as to why. 3 felt in front to be more distracting. 3 felt behind to be more distracting. 4 did not notice at all. Reasons given for not noticing were focusing on the task, just didn't know and commented that they were anxious after their absence (but not enough to look for them).

When laughter was heard and the actors were not in front of you, were you able to tell that the sound was coming from behind you?

This question was asked to confirm whether the spatial audio configuration was correct. If the subject could not tell the sound was behind them then it could explain why they didn't turn around if they didn't. 1 user didn't consider it at all, 6 could tell it was behind and 3 noticed laughter but didn't place it behind them. Additionally, comments were made on the fidelity of the spatial audio. Generally those who felt they heard it behind them also said it felt directional rather than spatial (direction and distance).

General Comments

Subjects commented mainly on design decisions or small moments of confusion with those decisions. Three commented on attempting to determine the pattern that triggered the actors' interactions. one user felt it was too long, which was due to an error which included 5 extra gestures (1st participant, resolved afterwards). two users commented on confusion with the ball system. One commented that the balls' colours, the red rest ones, didn't communicate consistently and clearly. This is due to a bug which occasionally causes one to flash pink when held. The behaviour of overlapping balls was not clear to one user. Additional impressions were that the overall experience was 'eerie' and 'weird'.

Chapter 7

Evaluation

This chapter evaluates the data collected throughout the experiment and considers its relevance to the questions raised and hypotheses and discuss the possible reasons for the results.

7.1 Completion Time

7.1.1 By Level

In 6.1 it is clear that certain levels take more time on average to complete. Since 2-side was designed as the baseline, containing no actor interactions, levels containing interactions should take longer. The evidence suggest this to be true , and a trend is demonstrated in the length of extra time required to complete a given level on average.

Three possible results were hypothesised as likely results. Either the actors from behind, which could significantly distract the subject, or the actors from the front, in which the user was likely to see interactions during their gestures, would be the most distracting.

Finally, the possibility of no result was considered. The actor interactions and presence may not modify the behaviour of the subject and their presence be almost entirely ignored by the subject.

On average, 7 front required the most extra time, followed by 2 front. 7 back, 2 back, and 2 side all requiring roughly the same amount of time.

With the frontal iterations, users were guaranteed to have the actors in sight, as well as their interactions. This was not guaranteed for the behind iterations, in which often the subjects did not turn around for external factors which will be discussed. Therefore, the question of whether in front or behind would disconcert the user more could not be answered from the results.

7.1.2 By interaction

Figure 7.1: Summary of average time for tasks of a given interaction

Back Rooms

Actor Interaction	Avg Time (s)
None	11.78
LookPlayer	11.78
LaughAtPlayer	12.02
SemiCircle	12.07
Laugh	12.81

Front Rooms

Actor Interaction	Avg Time (s)
None	11.99
Laugh	12.54
LaughAtPlayer	12.99
LookPlayer	14.10
SemiCircle	14.46

From the data tables were generated that express the average completion time for tasks containing a given actor interaction alongside the times without actor interactions for comparison. What was found was that interactions involving actors from the front caused the subject to take longer than from behind. This is shown in the 2 tables of which one contains averages for tasks with actors behind and the other averages from the front. Since often the subject did not notice the actors behind them the similarity between results with visual interactions behind and results without interactions is unsurprising.

The results in rooms with actors in front present a more interesting trend. In these instances it is known that the subject is able to see and/or hear the interactions. When the interactions are performed, the actors are within the line of sight of the subject. These actions do affect the average completion time, and to differing degrees. The most significant impact comes from the SemiCircle interation which involves the actors surrounding the user in a loose semi circle and staring at the user. This interaction caused some users to literally step back. Users reported feeling unnerved by the stare and this reflected in their behaviour by interrupting their gesture to look at the actors. This indicates some level of presence is achieved in a stimulus less intense than those used in other experiments such as the pit room or bridge, a question raised by Meehan et. al.[10].

7.1.3 Standard Deviations

Figure 7.2: Standard deviations for task completion by discrete task type

	Std Devs
ActionA	0.6628401184
ActionB	1.854175402
ActionC	0.3730753182
ActionD	1.225007217
ActionE	1.086358654
ActionF	1.246760256
ActionG	2.04215216
ActionH	0.9056175849
Cheerleader	1.360139677
Cross	0.7435501722
Hadoken	1.05507712
Hidden Dragon	0.93331682
Lefty Lunge	2.102592363
Lunge	1.876038769
Paint The Fence	0.6603632234
Reverse Starma	0.133572471
Starman	0.2199809196
Grand Total	1.45640219

The standard deviations of the individual actions indicate that certain tasks are able to be completed consistently and that some others had varying results. The greatest deviation was within the two 'lunge' activities. For most people, performing the lunge activities requires a step forwards on one foot due to the distance of the hands. Some subjects hesitated in moving their feet from their established position and attempted to reach the further ball from a standing position. In spatial terms, it would be obvious that on could not reach the further ball in a real-world scenario. That this is not obvious in VR indicates that spatial awareness might be affected in some way within the VE.

Other actions which presented more difficulty involved some form of physical movement. A few actions required the subject to kneel slightly to reach the balls closer to the floor. A crossover action where the left hand's ball and right hand's ball were on the right and left of the user respectively was shown to be more difficult

7.2 Qualitative response

A goal of this experiment was to identify whether the simulated actors, with animations, could produce expected responses to their actions.

From the interview many users elaborated on their attitudes towards the actors. I do not suggest that any of the subjects believed them to be real people, however their responses to the behaviours often matched expectations as if they were. The kismet experiment[2] and the use of eye images[5] to reinforce socially conscious behaviour suggest that an actor does not necessarily have to be believed to be real for people to be affected by their behaviour. The comments of unease when the actors stared at them and verbal comments during the experiment suggest that the relatively low-fidelity characters with simple behaviour and animations can produce genuine reactions.

7.3 VR Sickness

None of the 10 users experienced VR sickness. Reportedly a tolerance can be attained through frequent use. Since only 1 participant had over 2 hours' time with the Vive/Rift, the 9 remaining users could not have built a tolerance to VR. The steps taken to minimise the risk to an acceptable degree appear to have worked. IPD was measured, performance retained a stable 90fps, and users could only move within the environment by physically walking. Unfortunately, many evaluations cannot abide these guidelines, particularly the inability to move the user's position within the VE.

If the evaluation being performed allows for these requirements to be met then the results suggest they are likely to minimise VR sickness.

7.4 Turning around (or lack thereof)

The expected reaction when performing some task and unexpectedly one hears laughter behind oneself is to turn around and locate the source. Alternatively, they recognise that there are likely people behind them and continue regardess, which was the case for one user.

When performing the evaluation, many did not turn around to locate the laughter. The interview questions attempted to find the reasons for this. One possible reason a subject may not turn around is that they do not care to know. This may happen late in the experiment particularly if they are bored and just want to complete the task. One user claimed they intentionally directed their focus towards the task to maximise performance, and claimed that they would likely do the same in a realistic scenario.

I believe that steps can be taken to reduce this possiblity in the future. With the evaluation conducted, users were instructed to stand in a particular spot and have their attention directed in a specific direction. They were never asked to turn around during the experiment. Many of the users, most of whom were unfamiliar with VR, may treat the experience as similar to passive forms of media such as TV, movies or non-VR video games. In these scenarios one is not expected to physically explore, instead control the character through the direct use of the controllers or just sit and watch the screen.

If the experiment were conducted in such a way that the subject was required to explore the room in the tutorial, they may be more inclined to do so during the evaluation.

7.5 Limitations in Evaluation

There are a great many physical restraints while inside the VE. The user is not entirely certain where the physical walls are. They may be conscious of the possibility of hitting them and prefer not to move. While I only ever entered the playspace to assist with the equipment, the subject might not want to freely move in case they hit someone they can't see with the controllers.

To engage with the VE, the subject is required to:

- (optionally) put the wrist straps on from the controller
- wear a roughly 600 gram HMD with significant forward weighting
- adjust a rigid strap

- deal with a cable running down their back, making sure not to trip on it
- wear over-the-ear headphones

The constant reminders of the physical world and the equipment detract from presence. In addition to this, the Vive has an FoV of just 110 degrees at most, limited resolution and the limitations of the fresnel lenses can often be seen in high-contrast situations. Most of these observations were made during the evaluation in comments either during the interview, adjusting the equipment or during the experiment itself.

For these reasons I believe that the hardware involved in the evaluation actively discourage turning around, particularly when not familiar with the equipment. There is a risk of getting tangled. The limited FoV and limited area of sharp focus on the lens requires the user's head to turn to look around, instead of the natural method of looking around with minimal head movement. The considerable weight of the system makes it feel unwieldy and uncomfortable to turn. The headphones combined with the headset can not only feel warm particularly during some exercise, they cover a significant proportion of the user's face, potentially causing them to feel trapped or mildly claustrophobic. This is particularly problematic in experiments with measured stressful responses such as this one. The sensation of feeling trapped, unable to turn or move away from the actors could introduce additional stress beyond the expected level.

7.6 Optical Clarity

While efforts were made to ensure a clear image was attained by asking users to assess the clarity of the display's image, the subjectivity of such a request may contribute to a negative experience. Users were asked to read text and report discomfort, however many had never used a VR headset before. They had no reference for expectations of optical clarity. When I asked the subject whether the image was clear, they either said something was obviously wrong and the hardware was reconfigured, or they said it was ok and the experiment continued. This method catches the most obvious issues with image clarity such as badly configured IPD or to take glasses on or off. The distance from the frame of the goggles to the lens was never adjusted, which primarily is used as an FoV adjustment and to fit larger glasses. The main issue lies with the less obvious. Users without a reference for the expected level of visual fidelity for the given technology may recieve a very different experience than others if they cannot identify it. This could have an effect in instances where detailed visual elements appear in an experiment and the user may not be able to see it clearly. One user stated they had a lazy eye, an issue for which the vive has no ability to correct for (lens horizontal offsets are linked between each other, not adjusted independently by eye.). With the vive's lenses and a small 'sweet spot' in the center for which the image is clearest, this user could have had blurry vision in one eye. Issues with optical clarity may easily go unnoticed without more objective methods of determining whether the headset is optimally configured for the particular user. The process for determining clarity could be expanded on to include, for example, a typical eyesight test outside of the experiment followed by the same test within the VE.

7.7 Future Technology

VR as a consumer product is currently effectively in its infancy. Nintendo's Virtual Boy is so far removed from what modern HMD's offer that they are only technically of the same technology. With a fast moving field in largely untested waters, rapid developments to the issues identified with this first generation are to be expected. Many products which are being marketed as upgrades could potentially address some of the issues identified with the technology by the experiment.

A wireless Vive attachment, the TPCast[7], has entered the consumer market. This eliminates the need for the cable down the back and affords the user more freedom to move without the reminder of the cable. It runs for a few hours on a single charge and both powers the HMD and handles communication with the PC without major loss of visual signal fidelity. Currently stepping on the cable is likely to cause it to disconnect from the PC. This can cause massive disruption to an experiment.

Similar to my own solution, a rigid strap with integrated headphones will soon be available as an upgrade. My solution reduces the sagging feeling and the sensation that the HMD can fall off when sharply moved. The official strap should offer the same benefit. Alternatives to the Vive using the same technology except are lighter have already been announced e.g. LG's Steam headset[1]

The Vive's resolution and FoV were deliberate choices on Valve's part which were mindful of the graphical hardware constraints. The minimum required video card, a GTX 970, was considered a reasonably powerful card at the time. Increasing the resolution of the HMD would not have been sensible from a business standpoint as the price of the hardware to drive it would be too high. As graphics cards increase in rendering capabilities, HMD manufacturers can afford to include higher resolution, wider FoV screens in a smaller, lighter package. These smaller, lighter and higher resolution headsets can make the system more comfortable and natural for the user to work with.

Foveated rendering, a design which takes advantage of eye tracking in order to determine where the user is looking within the HMD, provides a shortcut in enhancing optical clarity[11]. If the center of the user's sight is known, then the area of the screen the fovea (Small region in center of eyesight) is concerned with can be rendered in high quality. With peripheral vision, primarily sensitive to movement, low quality rendering will likely not be noticed to the eye. With respect to this experiment foveated rendering could allow the user to freely look around instead of being constrained to the center of the lens. The integrated eye tracking could be used to gather additional information particularly concerning when they look at the actors. The question of whether the stare interaction causes the user to look at their eyes could be investigated with this. This information couldn't be gathered from just the video feed of the eyes, and if it were to be asked in an interview the subject would likely not be able to accurately state whether this was true.

7.7.1 Spatial Audio

Another identified issue with the evaluation was that 4 subjects could not discern the direction of the laughter. There are a variety of possible reasons for this:

- Incorrect hardware configuration e.g. left earpiece on right ear or volume too low
- Subject has poor hearing
- Incorrectly implemented spatial audio (audio driver or Unity)
- Improper headphones with regards to their capability to output spatial audio

Regarding the possibility of mismatched earpieces or incorrect volume, users were asked to ensure the volume of the tutorial sounds were suitably loud. Adjustments were made on a per-user basis. The older users preferred a higher volume.. Sennheiser HD598's are of an open-backed construction (which do not stifle the spatial characteristics) and are known for their capacity for spatial and directional audio, as well as clarity of sound without requiring an in-line amplifier, making them ideal for this particular situation.

The possiblity of the audio characteristics of Unity being the culprit seems likely. The default sound engines were used in the evaluation. The default sound engine is designed for stereo audio, but not specifically for headphones. The possibility of issues with the engine was noticed when Valve announced their Steam Audio SDK during the time of the evaluations.[23] Simulating binaural directional audio, the Steam Audio SDK appears to have been developed out of a need to improve spatial audio in VR. If their SDK provides what it claims to then I believe highly convincing audio can be introduced into VR evaluations. Of particular interest is the inclusion of spatial geometry in calculating sound. This could create situations where the user can differentiate between sounds around a corner and through a wall. The default audio options are likely insufficient for VR in terms of attaining presence and providing audio cues to users as evidenced by the number of subjects who could not discern direction.

Chapter 8

Conclusion

8.1 Summary

This dissertation's primary goal was to evaluate the use of new, low-cost technology in the context of the creation of experiments. Previous experiments involving earlier versions of VR technology were used as the inspiration for the evaluation. A suitable experiment was then designed and implemented in Unity. Running the experiment itself then revealed through both the qualitative and quantitative data the successes and limitations of the experiment in VR.

VR as a tool shows promise, with Users performing as they would be expected to in similar real-world situations for certain versions of the program. When the simulated actors interacted with the subject, their performance in the given tasks dropped as a result of their reaction to them. This effect was more pronounced with the larger crowd. In addition to this, certain tasks involving spatial elements were both reported as and measured to be more difficult than other tasks. The situations in which most subjects demonstrate behaviour that deviates from expectation demonstrate hardware limitations and characteristics that the design must consider in order to work around the issue. Iterations of the experiment involving elements behind the user generally failed to produce any response above the baseline. Designing an experiment involving elements behind the user require extra consideration, particularly with respect to the equipment's physical characteristics.

Based on the positive results when the actors were in front of the user, I believe that with some refinement and a carefully designed virtual environment the new VR systems can be used in the future as a low-cost and low-risk method of conducting an experiment.

Chapter 9

Future Work

Since this work deals with the first generation of technology of its kind, there leaves plenty of opportunity for running the experiment again in the future with later iterations of the technology. Spatial audio is seeing developments that can improve audible cues. Ergonomic issues present in the first generation are soon to be greatly reduced by technologies such as wireless HMD's, higher resolution displays with higher FoV and lighter, more comfortable headsets. The ergonomics were identified to be an issue to some subjects via their responses and their reluctance to freely move can be partly attributed to this. A more complete tutorial that encourages the participants to move around the room and familiarise themselves with the equipment may be enough to solve the issue.

The evaluation considers timings on a reasonably granular level. Additional metrics such as GSR, heart rate and gaze tracking could be used to obtain further detail into the relative effects of the interactions. One interaction designed to be stressful may distract users due to visual elements but cause no increase in stress, which would not be measured by the current methodology. Participants, for example, reported feeling uncomfortable in certain situations such as when there was laughter behind them. Currently there is no objective measure to quantify this within the experiment's data.

To avoid issues with VR sickness, the experiment was limited to a strict set of guidelines, such as maintaining a fixed virtual space. With developments in the understanding of the factors involved with VR sickness experiments making use of this information to expand the virtual environment beyond static space should be investigated.

While the data indicates some amount of distraction as a result of the actions, the sample size of 10 users performing a roughly 8 minute evaluation fit into 5 levels does not provide a large data set when the data is considered on a small scale. A longer, more focused version of this experiment with a larger sample of users investigating the more effective interactions could be performed to validate the results.

The next logical step after identifying the characteristics of a VR evaluation would be to replicate a well-understood behavioural experiment involving an environment and comparing the results. A suggestion for future work would be to virtualise an experiment such as a public goods game with multiple participants and virtual avatars and to compare the data. A firm baseline understanding of VR evaluations grounded in realistic scenarios can then be extended to allow VR evaluations to be conducted in unrealistic scenarios.

Appendices

.1 Attributions of sounds used

Here listed are attributions for the purposes of the Creative Commons License of the sounds used in the experiment:

- https://www.freesound.org/people/DanielsonIII/sounds/68399/
- https://www.freesound.org/people/DanielsonIII/sounds/68398/
- http://soundbible.com/1705-Click2.html
- https://www.freesound.org/people/jbeetle/sounds/274510/
- https://www.freesound.org/people/Ndheger/sounds/118116/
- https://www.freesound.org/people/fmaudio/sounds/152912/
- https://www.freesound.org/people/fmaudio/sounds/152911/
- https://www.freesound.org/people/fmaudio/sounds/152904/
- https://www.freesound.org/people/Reitanna/sounds/343934/
- https://www.freesound.org/people/Ch0cchi/sounds/15291/
- https://www.freesound.org/people/YvesIV/sounds/249718/
- https://www.freesound.org/people/chrillz3r/sounds/336534/
- https://www.freesound.org/people/Belloni/sounds/39862/
- https://www.freesound.org/people/Belloni/sounds/39863/

Bibliography

[1] ArsTechnica. Lg steamvr headset prototype. https://arstechnica.co.uk/gaming/2017/03/lg-steamvr-headset-prototype/. Accessed: 28/2/2017.

[2] Burnham, T. C., and Hare, B. Engineering human cooperation. *Human Nature 18*, 2 (2007), 88–108.

[3] Coburn, S., Rebenitsch, L., and Owen, C. Passive viewpoints in a collaborative immersive environment. In *International Conference on Virtual, Augmented and Mixed Reality* (2013), Springer, pp. 3–12.

[4] Cruz-Neira, C., Sandin, D. J., DeFanti, T. A., Kenyon, R. V., and Hart, J. C. The cave: Audio visual experience automatic virtual environment. *Commun. ACM 35*, 6 (June 1992), 64–72.

[5] Ernest-Jones, M., Nettle, D., and Bateson, M. Effects of eye images on everyday cooperative behavior: a field experiment. *Evolution and Human Behavior 32*, 3 (2016/10/18 XXXX), 172–178.

[6] Heilig, M. Stereoscopic-television apparatus for individual use. Patent application US2955156, Feburary 1960.

[7] HTC. Vive store page for tpcast wireless addon. https://www.vive.com/cn/accessory/tpcast/. Accessed 17/3/2017 Site in Chinese, TPCast unavailable on english versions of site.

[8] Lien, H.-C., Sun, W. M., Chen, Y.-H., Kim, H., Hasler, W., and Owyang, C. Effects of ginger on motion sickness and gastric slow-wave dysrhythmias induced by circular vection. *American Journal of Physiology-Gastrointestinal and Liver Physiology 284*, 3 (2003), G481–G489.

[9] Mark McGill, Alexander Ng, S. B. I am the passenger: How visual motion cues can influence sickness for in-car vr. In *Chi Conference for HCI 2017* (2017).

[10] Meehan, M., Insko, B., Whitton, M., and Brooks Jr, F. P. Physiological measures of presence in stressful virtual environments. *ACM Transactions on Graphics (TOG) 21*, 3 (2002), 645–652.

[11] nVidia. Innovative rendering technique that improves vr. https://blogs.nvidia.com/blog/2016/07/21/rendering-foveated-vr/. Accessed:28/2/2017.

[12] Oculus. Oculus best practices. https://developer3.oculus.com/documentation/intro-vr/latest/concepts/bp_intro/. Accessed:28/2/2017.

[13] okreylos. Optical properties of current vr hmds. http://doc-ok.org/?p=1414. Accessed 19/3/2017.

[14] Peperkorn, H. M., Diemer, J. E., Alpers, G. W., and Mühlberger, A. Representation of patients' hand modulates fear reactions of patients with spider phobia in virtual reality. *Frontiers in psychology 7* (2016).

[15] Phillips, L., Interrante, V., Kaeding, M., Ries, B., and Anderson, L. Correlations between physiological response, gait, personality, and presence in immersive virtual environments. *Presence: Teleoper. Virtual Environ. 21*, 2 (Apr. 2012), 119–141.

[16] PIMAX. pimax 4k vr headset. `http://en.pimaxvr.com/`. Accessed: 19/3/2017.

[17] RAMSAY, D. 3d virtual reality environment developed at uc san diego helps scientists innovate. `http://ucsdnews.ucsd.edu/archive/newsrel/general/09-083DVirtualReality.asp`. Accessed: 12-3-2017.

[18] RENNER, R. S., VELICHKOVSKY, B. M., AND HELMERT, J. R. The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR) 46*, 2 (2013), 23.

[19] SANCHEZ-VIVES, M. V., AND SLATER, M. From presence to consciousness through virtual reality. *Nature Reviews Neuroscience 6*, 4 (2005), 332–339.

[20] SEAN KELLEHER, S. Obtaining more wireless usb receivers. https://steamcommunity.com/app/507090/discussions/0/312265473878773762/. Accessed 19/3/2017.

[21] SUTHERLAND, I. E. The ultimate display. In *Proceedings of the IFIP Congress* (1965), pp. 506–508.

[22] TAFFOU, M., ONDŘEJ, J., O'SULLIVAN, C., WARUSFEL, O., DUBAL, S., AND VIAUD-DELMON, I. Auditory-visual virtual environment for the treatment of fear of crowds. In *Proceedings of the 2015 Virtual Reality International Conference* (New York, NY, USA, 2015), VRIC '15, ACM, pp. 2:1–2:3.

[23] VALVE. Steam audio. `https://valvesoftware.github.io/steam-audio/`. Accessed: 28/2/2017.