3D Perception of Complex Naturalistic Scenes in Consumer Virtual Reality

Rebecca L. Hornsey
Department of Psychology
University of Essex

A thesis submitted for the degree of
Doctor of Philosophy

June 2022

"keep your hopes up high and your head down low"

- A Day to Remember

Acknowledgements

Firstly, I would like to thank my supervisor Paul Hibbard. Thank you for your constant support over the years and for your positive outlook. I truly value you taking the time to not only teach me how to conduct proper research, but also for being a genuinely cool supervisor.

I'm grateful for the support my friends and family have given me during this time, particularly my mother Lindsay. This time has been made easier by your understanding and care.

Thank you to everyone inside the office and out, who gave up their time just to chat, run through experiments, and for the much needed coffee breaks. A special thanks to Jake and Jordi.

I am appreciative for the feedback my second supervisor Loes van Dam has given me during the planning for many of these chapters, and also for the Psychology Administrative Staff who were so pleasant to work with.

I gratefully acknowledge the scholarship from the Economic and Social Research Council which allowed me to carry out this research.

I have enjoyed this whole PhD because of you all.

Abstract

Vision research has progressed from simple experimental setups, to the immersive capabilities of virtual reality. This thesis will first introduce the reader to key concepts and background literature used throughout the text; a methodological chapter then outlines the practice and reasoning behind how software can be created for experimental studies in virtual reality. The first experimental chapter assesses 3D perception is two virtual reality headsets. Underestimation of distance in simple visual environments was found. Following this up, the second experimental chapter focusses on overcoming this issue using established methods: separating visual cues to determine which are most valuable and worthy of the host PC's processing power. Variability in distance estimation was found to be reduced by the addition of binocular cues, although an environment utilising all pictorial cues produced an equal sense of distance as binocular disparity alone. The third experimental chapter tests of the influence lighting, shadows, and clutter have on perception of virtual surroundings. Here it is shown that additional reference objects within an environment do not enhance distance perception, and visible cast shadows enhanced the precision of estimates. The penultimate chapter, an exploratory online experiment, is then presented which investigated perceptions surrounding immersive technologies. User experiences and impressions were investigated with both qualitative and quantitative data. Finally, a discussion on the results obtained through each stage of research, implications of these findings, and suggestions for future research is presented. Continuous advances in technology cause for continuous adjustments to be made, not only to the type of research being conducted, but also the type of equipment being used. This thesis supports the use of virtual reality as a means of gathering data in scientific research investigation, and also identifies specific visual cues that are vital in creating a realistic sense of depth within virtual scenes.

Contents

1 General Introduction **1**

1.1 Distance and Depth Perception . . . 2
1.2 Elements of 3D Perception . . . 3
1.2.1 Biological Cues . . . 4
1.2.2 Environmental Cues . . . 9
1.2.3 How Cues are Combined . . . 13
1.3 Perceptual Biases . . . 18
1.3.1 Cue Conflicts . . . 18
1.3.2 Distance Mis-estimation . . . 19
1.4 Virtual Stimuli and Display Devices in Vision Research . . . 23
1.4.1 Non-Immersive 2D Display Systems . . . 24
1.4.2 Non-Immersive 3D Display Systems . . . 25
1.4.3 Immersive Environments . . . 27
1.5 Development of Virtual Reality . . . 28
1.5.1 Requirements for Effective Virtual Reality . . . 29
1.5.2 First Attempts of Virtual Reality . . . 35
1.5.3 Current Virtual Reality Models . . . 37
1.5.4 Physical Space and Virtual Reality . . . 39
1.6 Inconspicuous Applications of Virtual Reality . . . 42
1.6.1 Distraction from Pain . . . 42
1.6.2 Educational Training . . . 43
1.7 Attitudes Towards New Technology . . . 45
1.8 Presented Research . . . 47
1.8.1 Power Calculations . . . 47
1.8.2 Summary of Chapters . . . 48

2 Using Unreal Engine for Creating Quantitative Experiments **50**

2.1 Introduction . . . 51
2.1.1 General Experimental Set-up . . . 51
2.1.2 Virtual Reality in Vision Research . . . 53

2.1.3 What Unreal Engine Is . . . 54
2.1.4 Applications of Unreal Engine . . . 56
2.1.5 Key for This Guide . . . 58
2.2 Creating an Environment . . . 59
2.2.1 Starting a Project . . . 59
2.2.2 Obtaining Assets . . . 64
2.2.3 Positioning Assets . . . 67
2.2.4 Asset Materials . . . 69
2.2.5 Lighting . . . 70
2.3 Stimuli . . . 73
2.3.1 Sound Effects . . . 73
2.3.2 Moving Objects . . . 73
2.3.3 Spawning New Objects . . . 75
2.3.4 Changing Appearances . . . 77
2.4 Timings . . . 79
2.4.1 Using Variables to Keep Track . . . 79
2.4.2 Events and Trials . . . 81
2.4.3 Creating Timelines . . . 82
2.4.4 Keeping Track of Time . . . 84
2.4.5 Setting Time . . . 86
2.5 Conditions . . . 87
2.5.1 Sequential . . . 87
2.5.2 Order Randomisation . . . 88
2.5.3 Counterbalancing . . . 88
2.6 Participant Interaction . . . 89
2.6.1 Controller Avatars . . . 89
2.6.2 Virtual Responses . . . 92
2.6.3 Environment Collision . . . 94
2.7 Data Collection . . . 96
2.7.1 Obtaining Raw Data . . . 96
2.7.2 Thresholds . . . 98
2.7.3 Saving Data to File . . . 99
2.8 Odds and Ends . . . 102
2.9 Exporting the Application . . . 108
2.10 Conclusions . . . 109

3 Size and Shape Constancy in Consumer Virtual Reality 110
3.1 Introduction . . . 111
3.1.1 Distance Perception . . . 111
3.1.2 Applications of Consumer Virtual Reality . . . 111
3.1.3 Biases in Distance Perception . . . 113
3.1.4 Shape and Size Constancy . . . 118
3.1.5 Current Study . . . 121
3.2 Methods . . . 122
3.2.1 Participants . . . 122
3.2.2 Apparatus . . . 122
3.2.3 Stimuli . . . 123
3.2.4 Task and Procedure . . . 124
3.3 Results . . . 125
3.3.1 Analysis Strategy . . . 125
3.3.2 Measurement Settings . . . 125
3.3.3 Effective Distance . . . 127
3.3.4 Correlation Between Size and Depth . . . 128
3.4 Discussion . . . 128
3.4.1 Summary of Findings . . . 129
3.4.2 Comparing to Prior Research . . . 130
3.4.3 Future Research and Conclusion . . . 130

4 Contribution of Differing Visual Cues on Distance Perception in Virtual Reality 135
4.1 Introduction . . . 136
4.1.1 Factors Affecting Distance Perception . . . 136
4.1.2 Methodological Set-up . . . 138
4.1.3 Current Study . . . 142
4.2 Experiment 1: Methods . . . 144
4.2.1 Participants . . . 144
4.2.2 Stimuli . . . 144
4.2.3 Apparatus . . . 146
4.2.4 Task and Procedure . . . 146
4.3 Experiment 1: Results . . . 147
4.3.1 Analysis Strategy . . . 147
4.3.2 Measures of Performance . . . 148
4.3.3 Summary . . . 149
4.4 Experiment 2: Methods . . . 150

4.4.1 Participants . . . 150
4.4.2 Stimuli and Apparatus . . . 150
4.4.3 Task and Procedure . . . 150
4.5 Experiment 2: Results . . . 152
4.5.1 Analysis Strategy . . . 152
4.5.2 Measures of Performance . . . 152
4.5.3 Summary . . . 155
4.6 Experiment 3: Methods . . . 155
4.6.1 Participants . . . 155
4.6.2 Stimuli and Apparatus . . . 156
4.6.3 Task and Procedure . . . 156
4.7 Experiment 3: Results . . . 156
4.7.1 Analysis Strategy . . . 156
4.7.2 Measures of Performance . . . 158
4.7.3 Summary . . . 159
4.8 Discussion . . . 160
4.8.1 Bias and Precision . . . 161
4.8.2 Comparing to the Hypotheses . . . 164
4.8.3 Interpretation of Findings . . . 165
4.8.4 Possible Explanations of Findings . . . 170
4.8.5 Conclusion and Impact . . . 173

5 The Role of Cast Shadows in the Perception of Distance 176
5.1 Introduction . . . 177
5.1.1 Environmental Context . . . 177
5.1.2 Shadows as a Visual Cue . . . 179
5.1.3 Naturally Occurring Shadow Confusion in Pictures . . . 184
5.1.4 Current Study . . . 186
5.2 Methods . . . 187
5.2.1 Participants . . . 187
5.2.2 Materials and Apparatus . . . 188
5.2.3 Shadow Conditions . . . 188
5.2.4 Task and Procedure . . . 190
5.3 Results . . . 191
5.3.1 Formatting Data . . . 191
5.3.2 Analysis Strategy . . . 192
5.3.3 Analyses . . . 194
5.4 Discussion . . . 197

5.4.1 Comparing to Hypotheses . . . 197
5.4.2 Interpretation of Findings . . . 200
5.4.3 Evaluation of the Methodology . . . 203
5.4.4 Future Research . . . 205
5.4.5 Conclusions . . . 207

6 Exploration of Perceptions Surrounding Virtual Reality Technology 212
6.1 Introduction . . . 213
6.1.1 Virtual Presence and Immersion . . . 213
6.1.2 Benefits and Applications of Virtual Reality . . . 214
6.1.3 Acceptance Models for Behaviour . . . 218
6.1.4 Acceptance Models for Virtual Reality . . . 221
6.1.5 Current Research Questions . . . 222
6.1.6 Mixed Methods Design . . . 224
6.2 Methods . . . 225
6.2.1 Setup and Procedure . . . 225
6.2.2 Data Collection . . . 225
6.2.3 Analysis Strategy . . . 226
6.3 Results and Discussion . . . 228
6.3.1 Background Profile . . . 228
6.3.2 Virtual Reality Experiences . . . 229
6.3.3 Concerns Associated with the Technology . . . 237
6.3.4 Opinions . . . 244
6.4 Conclusions . . . 249
6.4.1 Themes Found . . . 249
6.4.2 Implications for Manufacturers . . . 251
6.4.3 Comparison with Manufacturer Responses . . . 252
6.4.4 Summary . . . 255

7 General Discussion 256
7.1 Overall Findings . . . 257
7.1.1 Recap of Thesis Research . . . 257
7.1.2 The Value of Using Virtual Reality . . . 259
7.2 Future Directions . . . 263
7.2.1 Exploring the use of Unreal Engine 5 . . . 263
7.2.2 Higher-level Influences on Perception . . . 264
7.2.3 Multi-sensory Cue Combination . . . 265
7.2.4 Impact of Immersion on Perception . . . 267
7.3 Conclusions . . . 269

A Chapter 6 Questionnaire **271**

B Databases for Raw Data **276**

B.1 Size and Shape Constancy in Consumer Virtual Reality 276

B.2 Contribution of Differing Visual Cues on Distance Perception in Virtual Reality 276

B.3 The Role of Cast Shadows in the Perception of Distance 276

B.4 Exploration of Perceptions Surrounding Virtual Reality Technology 276

Bibliography **277**

Chapter 1

General Introduction

This thesis primarily explores 3D visual perception in virtual reality (VR), with a focus on depth and distance. VR is used as an experimental platform upon which to assess the accuracy of 3D perception, and the role of some binocular, perspective, and lighting cues in achieving this. The results of the experimental work presented throughout this thesis thus provide information on how the realism of the experience of 3D space within VR can be improved, and are relevant to the ways in which VR is designed for consumer applications. The empirical research presented also explores people's attitudes towards the use of VR, which are important to understand and accommodate if the potential benefits of VR across areas of application are to be realised. The core of the thesis comprises four experimental chapters and one methodological chapter. In this introductory chapter, an overview of the visual cues utilised during 3D perception will be introduced and key concepts used throughout the thesis will be explained. VR equipment will be introduced and its current applications discussed, as well as the proposition for its use in scientific research. This introduction will primarily focus on the optical factors that determine the experience of 3D space in VR.

1.1 Distance and Depth Perception

By depicting the visual trajectory of the eyes to a stimulus as a right-angle triangle, as shown in *Figures 1.1, 1.2, 1.3*, basic trigonometry can be used to explain the theory behind many different aspects of binocular distance and depth perception. It is not only the visual pathway which determines the way space is perceived, but also external factors pertaining to the object or environment, and the information that it therefore provides about distance and depth. Trigonometry can be used to calculate distance and depth from the convergence/viewing angle and the separation of the pupils of an observer. In *Figure 1.3*, the distance to the stimulus can be calculated if the distance between the observer's pupils (IPD) and convergence angle (θ) are known, using the *tan* function:

$$\mathtt{D} = \frac{\mathtt{IPD}}{\mathtt{tan}(\theta)} \tag{1.1}$$

The convergence angle can also be calculated by restructuring the equation:

$$\mathtt{tan}(\theta) = \frac{\mathtt{IPD}}{\mathtt{D}} \tag{1.2}$$

Consistent with the broader literature, *distance perception* is used here to refer to the internal estimate an observer has about the space separating themselves from a target. *Depth perception* refers to differences in the apparent distance between points on an object, and thus relates to the apparent 3D shape properties of a stimulus. In *Figure 1.4* two distances are depicted: $\mathtt{D}_1$ to the front of a stimulus and $\mathtt{D}_2$ to the back; the depth of this stimulus is $\mathtt{c}$, therefore depth is equal to the distance to the front minus the distance to back of the stimulus:

$$\mathtt{c} = \mathtt{D}_2 - \mathtt{D}_1 \tag{1.3}$$

Stereopsis refers to the vivid impression of solidity and depth, which can be achieved through the use of stereoscopic displays (Vishwanath, 2014; Vishwanath & Hibbard, 2013). Stereopsis is associated with both accuracy and reliability of depth perception

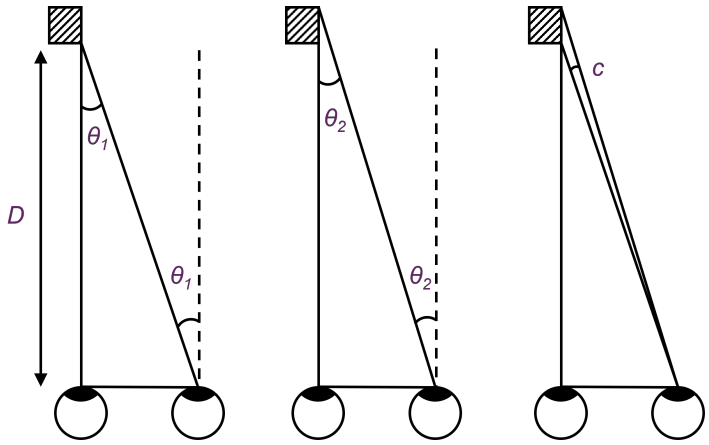


Figure 1.1: Two eyes observing a stimulus, where D is the distance to the stimulus; θ_1 is the convergence angle for the front of the stimulus; θ_2 is the convergence angle for the back of the stimulus; c is the difference in convergence angles between the front and back of the stimulus, and therefore perceived depth of the stimulus.

(Hibbard, 2008b) and it has been shown that the precision of this type of perception can be quantified through the use of 3D stereoscopic displays (Haines et al., 2015).

The human visual system utilises many types of information when interpreting depth and distance information, and weighs each cue in a manner that has shown to be effective through the millions of years of our evolution. Pictorial, monocular, and binocular cues are all important for accurate and reliable viewing: whether observing close-range or further afield distances, certain visual cues have an impact on the perception of one's surroundings. Many of the cues are connected, and so the theory behind how they work are intertwined also.

1.2 Elements of 3D Perception

Scene properties are aspects of a given scene which can provide information to the observer that is not limited to depth or distance information. Contextual information, colour schemes, and lighting are all examples of scene properties, and depth cues are as follows. There is overlap between depth cues and scene properties; *pictorial cues* are aspects of an image of a scene which include colour, texture, and lighting. These are also

examples of *monocular cues*, pieces of information from the scene which can be collected through one vantage point (or eye). Monocular cues are not limited to pictorial cues, but include motion parallax, perspective, familiar size, and accommodation. Additionally there are *binocular cues*, which make use of the visual components that are visible as a product of having two vantage points (two eyes), such as binocular disparity and vergence, which are discussed here.

1.2.1 Biological Cues

The distance between an individual's pupils is close to constant and therefore can be assumed to be intrinsically known; this IPD has a direct impact on the degree of vergence required for the eyes to binocularly fixate a stimulus. The variation in IPD for the general public can vary widely; according to Dodgson (2004) the mean and mode of adults IPD was found to be 63 mm, with a standard deviation of 4 mm. This is one point for consideration when analysing the experience of any head-mounted displays (HMDs) as it can result in altered perceptions of depth if set incorrectly. *Figure 1.2* depicts two situations where a differing IPD results in the perceived depth of an object to be different. The value of c_1 is the perceived depth which can be calculated using *1.3*. Due to the increased value of IPD_2 to IPD_1, the perceived depth of the object is greater in c_2 than in c_1.

When observing an object, two slightly different images are perceived by each eye due to the offset caused by the IPD. This difference in position on each retina is combined by the visual system and the differences in the two images contribute to the perception of distance and depth. The disparities within an observable scene range as a result of the degree of vergence (how crossed the eyes become when focused on a particular point), and the distance to the stimulus in question. Binocular disparity is the difference in location between the two eyes' images of any given point which is seen by both eyes, and is as a result of the offset of the eyes from one another. It is affected by viewing

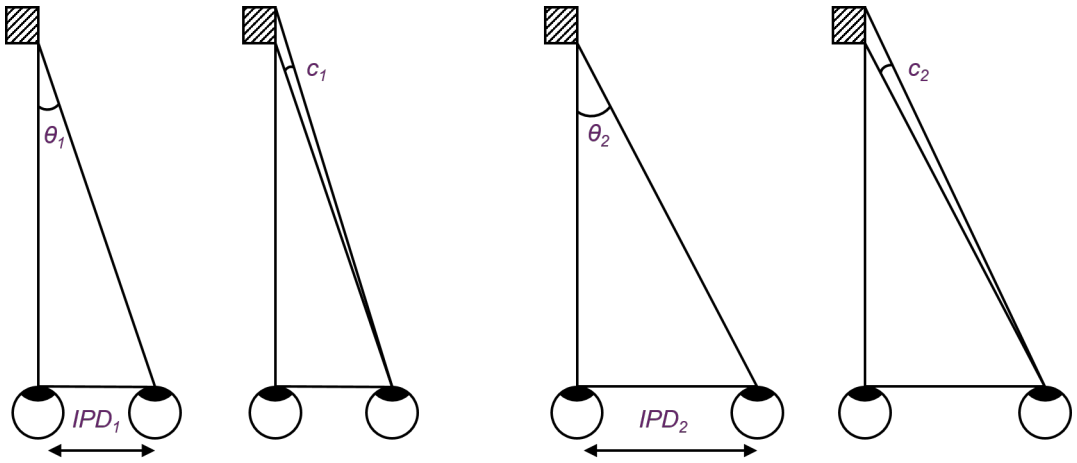


Figure 1.2: Two eyes observing a stimulus, where IPD_1 is a narrow distance between the eyes; IPD_2 is a wide distance between the eyes; θ are the convergence angles for the front of the stimuli; c are the difference in convergence angles between the front and back of the stimulus, and therefore the different perceived depth from a change in IPD.

distance as it is proportional to the inverse of the square of the viewing distance (Scarfe & Hibbard, 2006), denoting that an object will create a greater disparity when it is near than when it is far away. Depth discrimination from disparity has been found to be very accurate near the fixation distance but it worsens rapidly with increasing distance from fixation (Howard & Rogers, 2012) suggesting a natural need to incorporate multiple cues during natural viewing.

In natural viewing, both eyes tend to point to the same location when observing an object. When the stimulus is straight ahead in the distance, the directions of sight for the two eyes are effectively parallel to one other. The closer a stimulus is, the more each eye needs to rotate inwards to maintain observation on it: the amount of rotation is known as *(con)vergence* and is measured in rotation degrees deviating from straight, with 0° being the parallel position. *Figure 1.3* depicts vergence angles to two points on a stimulus: θ_1 and θ_2, when viewed from two distances D_1 and D_2. The convergence angle of the two eyes is a binocular cue which aids in determining absolute distance (Creem-Regehr et al., 2015). It has been suggested that convergence provides direct and unambiguous information about absolute egocentric distances to stimuli, which is due to unique convergence angles being established for near-distances (Richards & Miller,

1969), however its effectiveness diminishes with distance (Tresilian et al., 1999; Viguier et al., 2001). The eyes rotate in this way so as to allow the point of interest to be central in the fovea, where spatial resolution is highest; however if both eyes do not converge on the same object, potentially due to a weakness in one of the eyes, then double vision can occur. This sense of double vision, referred to as *diplopia*, means that the true location of an object in space, both horizontally and within depth, cannot be determined accurately through the use of this cue.

Contradictory to much of the literature on vergence, Linton (2020) suggests it is not an absolute distance cue, and the evidence stating otherwise has been produced in ways that are detached from everyday viewing, while also always in the presence of other confounding cues (often binocular disparity). A two-part experiment series was conducted by the author, the first of which tasked 12 participants to observe a collection of lights cast from a laser projector in a viewing box, and point to where they perceived the lights to be in exocentric space. The lights were projected at set distances within reaching distance (maximum 50 cm from viewing position). In order to identify the efficacy of the vergence cue, this was systematically manipulated in a step-wise manner while the focal distance was fixed through the use of lenses. The second used 12 undergraduate students as well as many of the participants from experiment 1. The initial set-up was replicated and the task was the same, however in this instance three lenses were used to elicit the focal response to be 'near', 'middle', and 'far' distances from the observer; the exact measurements of these distances was not given. The results overall indicate that participants were not able to accurately judge distance, when their vergence to a stimulus was gradually manipulated during this study. This does lead to a potential conclusion that vergence is not an effective cue in this situation, however the paradigm and set-up is far from what would be experienced in everyday viewing. This conclusion therefore cannot be applied beyond this environment.

Linton states that the current experiment is an outlier from other research which states this cue is useful because this is the only paradigm that has successfully divorced cues from one another. However, one source of information that is not removed or

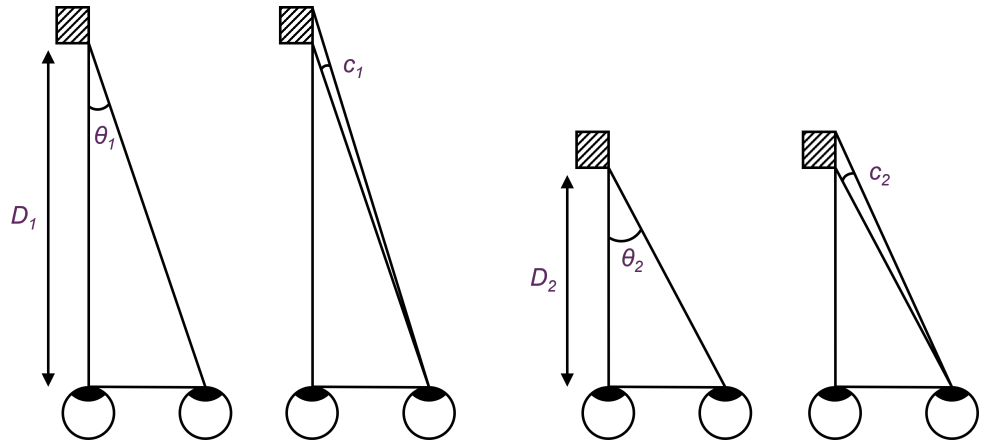


Figure 1.3: Two eyes observing a stimulus, where D_1 is a short distance between the eyes and stimulus; D_2 is a long distance to the stimulus; θ are the convergence angles for the front of the stimuli; c are the difference in convergence angles between the front and back of the stimulus, and therefore the different perceived depth from a change in viewing distance.

precisely controlled is the focus cues of accommodation and blur. Watt et al. (2012) built a multi-focal stereoscope which was able to present screen-based stimuli while having the capabilities of adjusting the focal and convergence cues.

Due to the step-wise variation of the convergence angle, the viewing conditions were not randomised and therefore subject to order effects. The order in which trials or conditions in an experiment are presented to participants can influence the responses given, which is why it is generally accepted to randomise or counterbalance the conditions experienced by participants. Here, not only were the conditions of the convergence angle and lenses not randomised, but they were gradually, and only slightly manipulated between conditions. It might be argued then than the result is an example of change blindness, in which changes that would otherwise be very salient go unnoticed if introduced very gradually Simons et al. (2000).

Another confound of the research is that in the first experiment there was a limited and potentially biased participant sample. The participants were recruited from the author's social media account, who arguably could not be considered to be naive to the research topic. It would not be inconceivable to assume the author's online friends may have at least some knowledge on the type of research he conducts and are not

representative of the general population.

Linton states that responding to the conscious awareness of eye movements is not visual processing, and that there is a distinction between simply *perceiving* and higher cognitive function of *attention*. Visual processing is a biological response to external stimuli and there are many physiological responses, from pupil dilation and lens thickness, to excited orientation columns, that all occur in response to visual stimuli. It is not clear whether and where a meaningful distinction can be made between those physiological responses that people are aware of (such as cortical excitation) that should count as perception, and those that should not. The author compares the results found to examples of change blindness, such as the expanding room (Glennerster et al., 2006) where observers do not notice changes in the scale of the virtual environment they are in; and colour change detection (Simons et al., 2000) where observers do not notice sections of an image changing colour over time. While easy to follow the argument for these being similar, registering a change in pictorial stimuli is not equivalent to registering a physical, internal change in the angle of eye rotation.

Accommodation is the change in shape of the lens to allow focus on a specific distance, controlled by the ciliary muscles in the eyes. This monocular cue is affected by focal distance but also the light of an environment: when in brighter environments the pupil will constrict to allow less light into the eye, and when in dimmer environments it will dilate to allow more light in. This has a knock-on effect for the shape of the lens, and when dilated the lens becomes more important in focussing. At near-distanced points of focus, the accommodative response is to constrict and thicken the lens to shorten its focal length, similar to how the ciliary muscles affects pupil size in the presence of bright light. When attempting to focus on an object at further distances, the lens adjusts to be flatter so that there is a decreased curvature for a lower refractive power. This is depicted in *Figure 1.5*, where the importance of correct lens power when refracting an image onto the retina are shown.

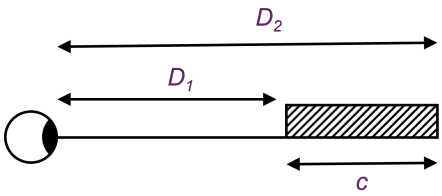


Figure 1.4: An eye observing an object, where D_1 is the viewing distance to the front of the stimulus; D_2 is the viewing distance to the back of the stimulus; c depth of the stimulus.

Retinal blur is another monocular focus cue, whereby stimuli at the distance of focus are at their sharpest perceivable state, and everything at closer or further distances will tend to be blurred. In natural viewing, for an object to have a sharp focus on the retina, the accommodation distance should be close to (if not, at) the object distance (Hoffman et al., 2008). Vishwanath and Blaser (2010) investigated the effect blur has on distance and depth perception by varying the degree of blur in artificial environments. Although this study investigated artificial blur, the conclusions made were applied to retinal blur and it was suggested that a blur gradient is likely to be a quantitative cue for distance perception, but would be most effective for viewing nearer scenes rather than far. Additionally, a study by Held et al. (2012) also investigated the role of blur in depth perception via a depth-discrimination task. Here, it was determined that blur provides greater depth precision than disparity. In *Figure 1.6* the effects of blur are presented, although these photos were taken by a camera, the effect is similar for retinal blur: objects which are being observed are in-focus, while surrounding objects are out of focus.

1.2.2 Environmental Cues

Characteristics from an environment and pictorial cues are also influential in the way 3D scenes are perceived. However, these are much simpler than the biological responses to viewing scenes as these can be used to portray depth in 2D images and do not require the use of stereoscopic vision to work.

When observing two parallel lines drawn across the saggital plane it would appear that they converge as distance increases, this phenomena is the result of the pictorial cue of linear perspective. So far as using this in the perception of depth, when observing the keyboard and edge of the table in *Figure 1.6*, the edges appear to slant inwards as they begin to meet the back wall. However if viewing the same scene from above, the edges would be parallel and void of this convergence. Testing how this effects distance perception, Thomas et al. (2002) used depth matching estimates and compared the results of conditions with vertical lines versus horizontal ones in the environment. It was found that the estimates were most accurate with linear perspective vertical lines. In addition to this, Surdick et al. (1997) examined perspective cues (with the inclusion of linear perspective and texture gradient) and found that these cues increased participants' accuracy of distance perception more than other cues tested, including texture and binocular disparity.

The effect of luminance on the perception of depth is reliant on a light source and its direction. Due to the occlusion of objects and shadows, as well as texture gradients, it tends to be the case that the further an object is away in a local surface region, the darker it will be. Shading on an object's surface can be used as an indicator of its shape, additionally the difference in shading can be an indicator of depth perception, where dark areas may suggest a further distance than lighter areas on a surface. Hibbard et al. (2017b) conducted an experiment to see if the manipulation of luminance affected the perception of depth within scenes in a global or local sense. In the global condition, a series of objects within a rendered scene were edited so that the closer objects increased in darkness, dependent on how far away from the centre of the scene they were, while those towards the back increased in lightness. A second global scenario had the same editing take place however enhanced the lightness of the closer parts and enhanced the darkness in the far. In the local scene, the same technique was used, however the individual pixels of each object were edited so that they were either lighter or darker than the mean luminance of the object. This set-up tested the notion that darkness means further

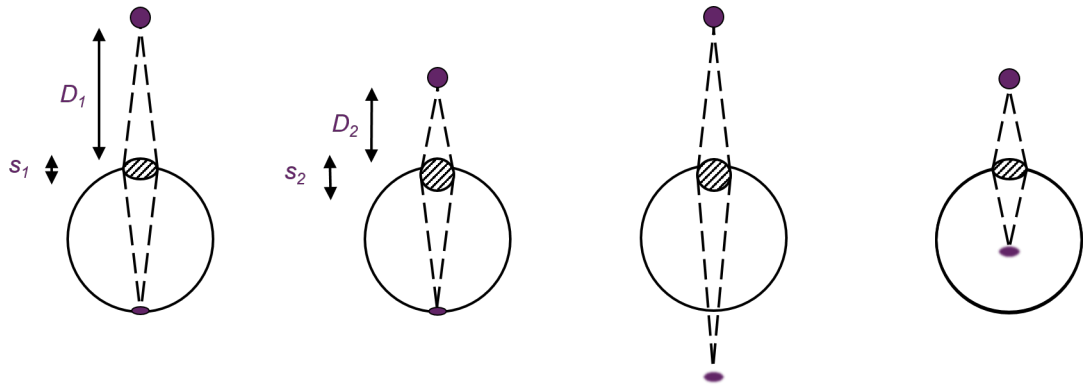


Figure 1.5: An eye observing a point in space, where D_1 is the viewing distance to a far point in space; s_1 is lens size for a point in far space; D_2 is the viewing distance to a point closer in space; s_2 is lens size for a point in near space. Here, the accommodative response allows for the point to be in focus, reflected onto the retina (from combinations D_1 and s_1; D_2 and s_2) or out of focus, with reflection beyond (D_1 and s_2) or shorter than (D_2 and s_1) the distance of the retina.

away than light areas. Participants adjusted an on-screen ruler in accordance with the amount of depth seen between two identified points, it was found that the luminance edits had no effect in the global condition, but in the local condition more depth was perceived where the natural luminance was enhanced. These results suggest that this pictorial cue is utilised significantly during scene observation and is an important feature for depth perception.

Shadows are dynamic in shape in the physical world and have been found to influence the perception of distance: in augmented reality (AR) for example, a study by Jurgens et al. (2006) found that when a shadow is cast directly under a stimulus, with a radius matching the distance from the ground, the accuracy of participants locating a target was not affected in a distance estimation task. One should be hesitant about generalising results from studies undertaken in AR to what might be expected in VR, because the addition of physical stimuli amongst virtual ones causes conflict in the perceived properties of the different objects (Kirkley, 2004). Using specific lighting in AR raises the issue of conflicting light sources: the lighting needs to be consistent with the lighting of the environment in which the user is physically in (Sugano et al., 2003). Cast shadows can also be used to determine direction and intensity of light source (Casati, 2004; Koenderink et al., 2003; Xia et al., 2014), as well as the shadow boundary being

found to impact the perceived depth between the stimulus and surface where the shadow is cast (Kersten et al., 1997).

When comparing the observable features of objects at differing distances, for example the keyboard in *Figure 1.6*, it is obvious that there is more texture information (such as individual keys) observable at closer distances than far, and it may be possible to use this as a way of determining a sense of distance (Bishop, 1989). Another example would be looking at a tiled floor it may be possible to count the number of tiles to one thing and know that double the number of tiles would be double the distance. With this in mind, an influential psychophysical experiment by Whitehouse and Gruber (1957) used a distance bisection task in real space to examine size and distance judgements as a function of texture: the procedure required participants to identify the mid-way point between themselves and a reference point within an environment which had either some or no texture cues. They found that the mid-way point was set further away than where it should have been, regardless of the texture condition experienced. The conclusion that there was an over-estimation of near space, from the over-setting of the half-way points, was made and that texture cues were not used; however, this generalisation perhaps should not have been made given that the texture presented in conditions was not systematically varied. From this, Wohlwill (1963) further investigated the role of texture in distance perception with another bisection task using conditions varying in texture itself and texture density. The author here was overt in their descriptions of texture: patterns of discrete elements with constant size and shape. In this experiment, stars on the floor of the observable environment were used as the reoccurring texture pattern. It was found that there was again this over-estimation of the mid-point in all conditions, even when void of texture, however the varying degrees of texture only affected the results minimally. More recent findings are still supportive of this; Witmer and Kline (1998) found the texture of a floor did not reliably affect distance estimates, but the addition of other external cues on top of texture did increase the accuracy to near-perfect.



Figure 1.6: Two photos of the same scene, with far (a) or near (b) distances in focus, causing different items to be blurred. Also showing perspective as a distance cue with edges of the keyboard and table converging as distance increases.

1.2.3 How Cues are Combined

The many visual cues which are available during observation are combined by the visual system to provide a perception of distance and depth. The precision that each cue provides, in terms of distance estimates, has been shown to vary across cues (Cutting & Vishton, 1995). One chapter within this thesis will examine the way different visual cues are combined. The optimal way for the visual system to combine the information from the cues is to determine which cues provide the most reliable estimates and which provide the least reliable estimates. The most commonly accepted theory is that the visual system uses all available information, but relies more heavily on cues which have a smaller degree of variation than the cues which produce large variations in estimates. This is often referred as MLE (maximum likelihood estimation) which is a form of an ideal observer model. This is generally accepted because it would be non-optimal for the visual system to combine cues in a fashion that does not produce the most precise perception of distance, and combining cues in this way optimises the precision of the resulting estimate (Oruç et al., 2003). This topic is experimentally explored in *Chapter 4*, while an overview is given here.

In natural viewing of a physical environment, we assume that two visual cues elicit a similar estimate of viewing distance to a surface. Because of the imprecision in the encoding and decoding of this information, it is likely that the two estimates will be slightly different; the visual system will combine the estimates using a weighted average, with the weights based on their standard deviations (Oruç et al., 2003). In *Figure 1.7* are the distance estimations of a given stimulus, and the likelihood of the distance that the visual system will interpret, for two types of cues (given here are pictorial and binocular cues). The estimates from binocular cues show less variation in the distances, as represented by a slimmer curve with a higher likelihood, compared to the larger spread of estimates from pictorial cues. Cochran, 1937 has shown mathematically that by combining the estimates from both cues, a more reliable estimate can be made, the weightings given to each cue should be inversely proportional to the variance. Because of this, reliability can be defined as $1/variance$, where a cue providing larger variation in estimates would be considered less reliable. There is evidence of this type of proportional combination from many studies and across many modalities (Ernst & Banks, 2002; Glennerster et al., 2006; Scarfe & Hibbard, 2011; Watt et al., 2005a).

Bayesian logic can be applied here, to understand how the likelihood functions from two cues can be combined. In order to calculate the optimal way these two cues can be combined, the two likelihood functions can be multiplied together, resulting in a third likelihood function with a mean that is positioned somewhere in the middle of the means from the initial cues, and has a more compact distribution than the individual distributions. The exact location of this new mean estimate will depend on the individual distributions, but will be shifted in the direction of the more reliable cue. This method is appropriate for any shaped curve of distribution without needing to assume a particular form of the likelihood, such as a Gaussian distribution.

If the distribution curves are however Gaussian, then the method is the same as the weighted averaging model of cue combination. The weighting allocated to each cue when estimating a distance is still determined by the reliability of these cues, in the form of

the standard deviations of the estimates. The difference here is that there is no point multiplication from the estimates between the cues, only the relationship between the means and standard deviations are taken into consideration when applying a weighted averaging approach. In the example of *Figure 1.7* the distance estimation with the highest likelihood is the same for both types of cues, so the distance estimation with both cues combined will be identical, but can be thought of as more reliable. If the two cues had conflicting distance estimates (curves shifted left or right along the X axis to alter the sigmas) then the combined estimate would lay somewhere between the two curves, with a bias in the direction of the more reliable cue.

According to these models, the weighting of each sensory cue is defined by its variability relative to other cues. Observers should be more accurate when multiple cues are available to provide depth information. A study by Hillis et al. (2004) tested this in a paradigm where only disparity or texture were visible to participants, the variability of surface slant estimates corroborated the model, and the conclusion that the visual system weighted the contribution of each cue according to its relative reliability. Evidence for this method of cue combination has been found not only from visual cues (Keefe & Watt, 2009; Scarfe & Hibbard, 2006), but also across sensory modalities (Alais & Burr, 2004; Ernst & Banks, 2002; Fetsch et al., 2012).

If two cues are to be combined accurately, it is important to establish that they relate to the same physical property. For example, two estimates of the distance to an object should be combined if we are confident that they refer to the same object, but not if they are likely to refer to different objects. Models such as the causal inference model have been developed to address this problem (Körding et al., 2007).

Increasing the number of cues available will have an impact on any biases and accuracy of distance estimates, however it is difficult to include every single cue that may be available in naturalistic viewing, due to the vast amount of possible cues. The less reliable the additional cues are, which are not included in any given model, the larger the impact they will have on the model predictions. In VR for example, the focal distance

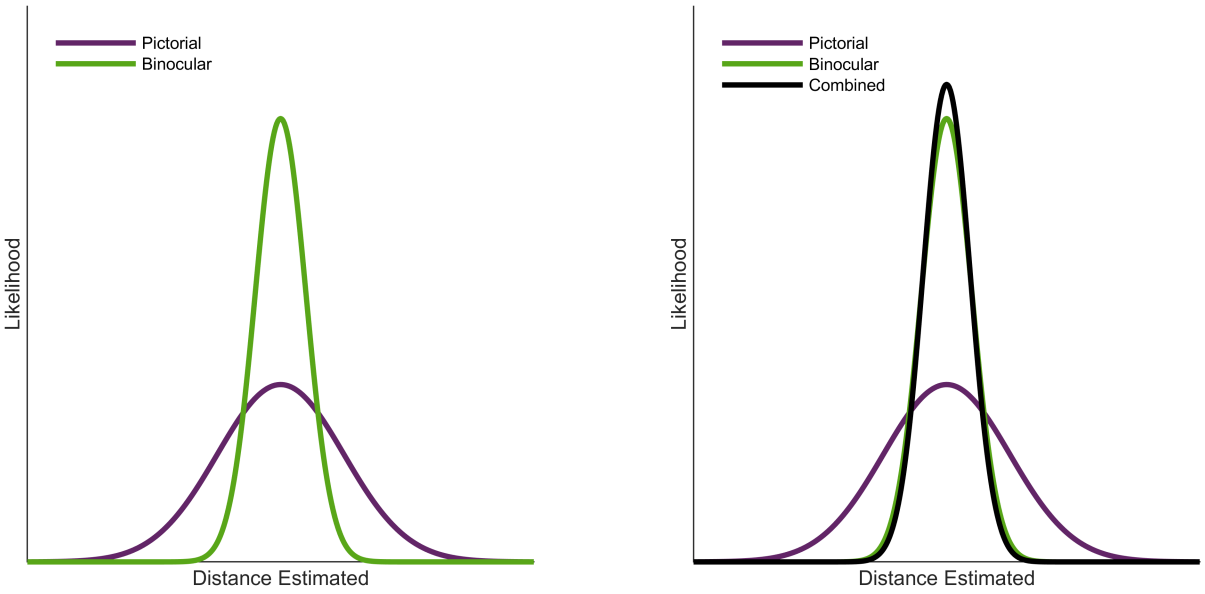


Figure 1.7: The way two cues, such as pictorial (purple) and binocular (green) cues, are combined in a weighted-manner (black line).

is often set by the manufacturers and unable to be adjusted accurately; these can have an influence on perception when using these devices, and may be ignored factors in cue combination models.

This cue combination model requires an estimate of the mean and variance of the likelihood function available from each available cue. This could be implemented neurally in a population of neurons, with tunings that vary along the dimension of the property to be estimated. The shape of this distribution of activity across the neural population could be used to provide an estimate of the mean and variance of the likelihood function that is required (Battaglia et al., 2003). This model is explored further in *Chapter 4*.

In contrast to these combination models, a cue veto model (Bülthoff & Mallot, 1988) has been proposed, in which any unreliable depth cues are ignored, and the estimate is based on the cue judged to be most reliable. This model instead proposes that the visual system rejects entirely the estimates from an unreliable distance cue. In this instance, the authors discussed how edge-based stereo produced a better depth perception, compared to that from shading or disparate shading, causing the information from these two shading cues to be ignored as opposed to combined.

Hartle and Wilcox (2021) came to the same conclusion, based on the result of their experiment where binocular disparity and motion parallax were tested independently in terms of their contribution to distance estimates. The authors suggested that the weighted combination model does not adequately describe performance in physical or virtual environments, as they tested the same cues in both a physical and virtual set-up. The authors concluded that motion parallax was less reliable than disparity, and when combined the distribution was not significantly different from the distribution of binocular disparity alone. Much less literature is published on this model compared to the combination model, and few studies have tested this with rigorous method (Arnold et al., 2019), however advancements in the technology used to test this may lead to different ways of comparing the models.

Instead of a different model, it could be argued that the veto model is the same as the weighted model. In this case, the most unreliable cue might be providing such a small contribution that it is difficult to determine whether this is the case or whether it is being vetoed. It would be hard to show the effect of such an unreliable cue, and huge amounts of data would need to be collected in order to make this distinction.

The reliability of cues are subject to change based on the viewing distance (Loomis & Philbeck, 1999; Nagata, 1991; Surdick et al., 1997). As shown in *Figure 1.3* the convergence angle gets smaller as distance to a stimulus increases, which reduces its effectiveness as a distance cue. This is the same for other biological cues, such as binocular disparity and accommodation, yet the effectiveness for environmental or pictorial cues has a greater range.

1.3 Perceptual Biases

1.3.1 Cue Conflicts

When two or more cues are visible, but signal different distances to a stimulus for example, this is known as *cue conflict*. A common conflict is found in conventional 3D displays where unnatural conflicts between vergence and accommodation are created, as these cues are usually coupled in natural viewing (Howard, Rogers et al., 1995). When viewing a stimulus via a stereoscopic computer monitor, the distance of the rendered object may appear to be closer than the screen itself. The convergence response would be to focus at the virtual (binocularly specified) distance, however as the two images being presented to each eye are actually being shown on the computer *monitor*, accommodation would be to this physical distance. This conflict between the physical vergence angle of the virtual stimulus and the accommodative response to the monitor may affect the ability to fuse binocularly, and the perception of distance.

With regard to VR headsets the screens are only a few centimetres away from each eye, and lenses are used to ensure that the images can be focused. These lenses shift the accommodative response to a fixed distance in near-space (Elias et al., 2019; Kramida, 2015): this is where the issue lays. Accommodating to objects at a close-range while converging to a mis-matched, further distance may be the reason for eye-strain when using HMDs like this (Hoffman et al., 2008). One way around this is the use of multi-focal displays: Mercier et al. (2017) developed a system which incorporated vergence and accommodation eye tracking and had three adjustable-focus displays per eye. This set-up was not appropriate for consumer use, but will likely make appearances in future VR equipment as the cue conflict reduction results in less unwanted by-products from the use of HMDs. There have been other attempts at multi-focal displays (Aghasi et al., 2021; Hu & Hua, 2014; Liu & Hua, 2010), however these set-ups are inherently bulky and created for very specific experimental laboratories.

Retinal blur from the accommodative response is an effective distance cue in physical

space (Held et al., 2012; Held et al., 2010; Vishwanath & Blaser, 2010), however in most of the current VR headsets, there is a fixed focal distance as previously discussed. Due to the fixed focal distance, when observing through the prisms of the headset, all objects within the environment will appear to be in-focus as the accommodative response interprets a singular distance. Therefore, retinal blur can no longer be used as a distance or depth cue, as there are no focal variations in distance within the environment, biasing everything to appear at the same distance. The equi-distance tendency, where multiple objects are judged to be at similar distances when presented in in sparse environments, is found in physical space (Gogel, 1965), this phenomenon may only be exacerbated by the fixed focal distance of current VR devices.

1.3.2 Distance Mis-estimation

The perception of distances has been measured in a number of different ways: direct methods include verbal estimates (Loomis & Philbeck, 2008; Loomis et al., 2002; Messing & Durgin, 2005), physically moving (Loomis, Knapp et al., 2003; Thompson et al., 2004; Witmer & Sadowski Jr, 1998), or positioning objects at a desired location (Ellis & Menges, 1998; Nakano et al., 2008; Ponto et al., 2013). Indirect methods gather estimates of another metric, from which the perceived distance can be inferred; prime examples of this are size and shape constancy tasks (Brenner & van Damme, 1999; Carlson, 1960; Scarfe & Hibbard, 2006). There is however, evidence to suggest the fidelity of responses differs depending on the task at hand (Jerome & Witmer, 2005; Napieralski et al., 2011). A study by Glennerster et al. (1996) compared the distance estimates of a stereoscopic surface when participants completed direct and indirect measures of distance judgement tasks. The results showed that accuracy was higher when a direct matching task was completed, rather than when participants were required to construct a metric representation in the indirect task. This suggests estimates are often task-dependent, although errors in judgements are consistently found. This is explored throughout the experiments presented in this thesis.

A reoccurring issue found in a lot of virtual environments is that distance is underestimated (Waller, 1999) and items presented within the virtual environments may not then be seen to have the intended shape or size. Loomis, Knapp et al. (2003) suggested that the reason for depth compression in immersive virtual environments was because the scene rendering was lacking in visual cues that, although subtle, had a major impact on distance perception. This included the likes of natural textures and highlighting; continuing with this logic, any scene that includes natural (or at least naturalistic) visual features such as these would lead to distance perception comparable to that found in real space. As discussed, VR technology is still improving, so this veridical experience is yet to be created.

Thompson et al. (2004) composed an experiment with this theory as the basis; a blind walking task was conducted with four conditions: (1) in real space; (2) a virtual wire-frame mesh; (3) a 360° high-resolution panoramic environment; (4) a low-quality texture map. An nVision Datavisor HiRes HMD was used for the three virtual conditions, this headset had a 42° horizontal FoV and a resolution comparable to some HMDs currently available with 1280 x 1024. The IPD was set to 65 mm which was consistent with the panorama images used as stimuli, and participants were allowed to observe each environment before the task took place. The results showed that in all simulated environments the amount of compression was relatively similar, and significantly more than that perceived in the real space condition. The conclusion drawn from these results was that the compression effect found within virtual environments was not an effect of the quality of graphics or rendering used. Thompson et al. suggested that a degraded sense of presence while wearing the HMD may instead have caused the mis-estimation. This interesting conclusion can then be used as the basis of future experiments to explore the effect graphics and other environmental cues have on the subjective experiences of these virtual environments, in terms of sense of presence and enjoyment.

Presentation distance of a stimulus elicits an altered perception of the depth within

the stimulus: *Figure 1.3* depicts the way in which adjusting the viewing distance from nearer to further reduces the depth perceived, as c_1 is smaller compared to c_2. Due to many internal cues being less effective at longer viewing distances (Kline & Witmer, 1996; Loomis & Philbeck, 1999; Nagata, 1991), the reason for viewing distance having such an impact on shape and size perception may be due to the combined lack of available information from these cues.

Constancy tasks are an indirect way of measuring distance perception: if an object is constant in its size or shape, then these properties should not change when presented in different viewing conditions (such as presentation distance). If a physical ball is moving through space, or is first shown at a close and then a far distance, with unchanging physical dimensions, then the perception of these dimensions should match. However, there is evidence for a mismatch of the perception of these qualities and the true values in both real space and for binocularly-viewed, computer-rendered stimuli. *Figure 1.8* depicts a constancy task in VR, where a virtual sphere is presented at a set distance in each trial, and the observer is tasked with adjusting the size or shape of it to match that of the physical reference they hold.

A slight rearrangement of *Equation 1.2* can be used to explain how a constancy task can be used to determine the distances an observer perceives. The rendered size of a stimulus ($\mathtt{h}_{\mathtt{rendered}}$) in a trial, and presentation distance ($\mathtt{D}_{\mathtt{rendered}}$), are assumed to be equal to the perceived size (size setting made in the trial; $\mathtt{h}_{\mathtt{perceived}}$) and perceived viewing distance from the observer ($\mathtt{D}_{\mathtt{perceived}}$):

$$\mathtt{tan}(\theta) = \frac{\mathtt{h}_{\mathtt{rendered}}}{\mathtt{D}_{\mathtt{rendered}}} = \frac{\mathtt{h}_{\mathtt{perceived}}}{\mathtt{D}_{\mathtt{perceived}}} \tag{1.4}$$

The assumption here is that what is being presented is actually what is being perceived. This can then be further rearranged to express the perceived distance from the perceived size:

$$\mathtt{D}_{\mathtt{perceived}} = \frac{\mathtt{D}_{\mathtt{rendered}}\mathtt{h}_{\mathtt{perceived}}}{\mathtt{h}_{\mathtt{rendered}}} \tag{1.5}$$

This transformation allows for the perceived distance to be calculated. This calculation assumes that accurate size scaling was achieved, and therefore any inconsistencies in the size judgements are a result of incorrect distance perception to the stimulus.

A study by Johnston (1991) tested this through the use of a random dot stereogram, with two participants completing 200 trial repetitions. Shape judgements were made by participants in that they were required to adjust the depth of a cylindrical stereogram until it appeared to be accurately cylindrical: not stretched and not squashed. It was found that when presentation distance was near, perceived shape was elongated; when presentation distance was far, perceived shape flattened. The authors concluded these errors were due to the incorrect measure of egocentric distance.

To continue with this methodology, a study by Brenner and van Damme (1999) found similar results when size and shape judgements of a virtual ellipsoid were made against a physical tennis ball reference. Judgements from six participants were made when the ellipsoid was presented at various distances; viewing was restricted by the use of a chin-rest and the amount of visual cues was moderated by having a dark or dimly lit experimental room. This experiment, and many others involving similar paradigms (Beusmans, 1998; Glennerster et al., 1996; Loomis et al., 1992) provide consistent evidence of how object size and shape are mis-estimated, by the way of over-estimations made at near distances, and under-estimations at far distances. This suggests humans do not exhibit shape constancy.

Further to this, a study by Scarfe and Hibbard (2006) examined whether seven participants' size and shape judgements for a stereoscopic cylinder were in accordance perceptually, when presented in motion. Judgements of these metrics were made at four egocentric distances, followed by a presentation of these estimates warping through the distances and estimates made. Results showed that the apparent structure of the stimulus was not consistent when presented at differing distances, and participants could not accurately identify constant shape when in motion. Further investigation into what is causing these errors are necessary; perhaps VR can provide some evidence and point in a suitable direction for the cause of this issue due to the unconstrained viewing, if

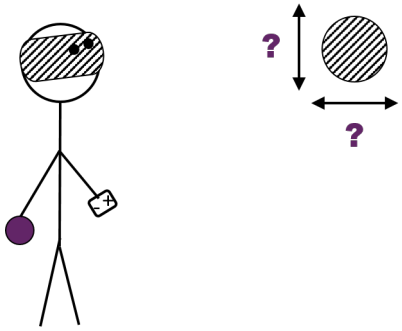


Figure 1.8: Constancy task, a participant views a stimulus and adjusts the size of it to match that of a physical reference. The size estimate can then be used to calculate the viewing distance that the observer perceived in that trial.

similar distortions even occur when using this technology.

1.4 Virtual Stimuli and Display Devices in Vision Research

The focus of this thesis is the use of head-mounted VR; a display which presents two separate images to the eyes of the observer to approximate the two monocular fields of view each eye would receive in natural viewing of the physical world. Controllers and tracking systems can be used in combination with these headsets, depending on whether the experience is to passively observe or actively engage with. Areas of research which come under the umbrella term of *visual perception* are very broad, as are the types of equipment that can be used when testing in this field. In order to understand how VR can be used in vision research, and how to optimise the perceptual experience in VR, it is helpful to discuss it in the context of other display-screen technologies. Explained here are a few examples of the types of equipment used in vision research, as well as the equipment used for this thesis, in addition to brief examples of ways their specifications can be utilised in this field.

The starting point for understanding vision is the ambient optic array, the structured arrangement of light with respect to a point of observation. As human observers have two eyes, a consequence is being able to sample the ambient array from two points,

defined by the centres of the eyes. The ambient optic array is then defined by the light arriving at these points from every visible direction, with each direction projecting to a different location on the retina. These retinal images then form the input stimulus for visual processing.

The ultimate display device for use in vision research, and in other applications, would be one that was perfectly able to replicate the visual input from a real-world scene. Such a display would then be expected to pass the 'computer graphics Turing test', such that an observer would be unable to tell the difference between the display and the real scene (Banks et al., 2016). In practice, and in the current state of the art for consumer VR and other displays, the display will be limited by hardware and software reasons from exactly replicating the visual input from a natural scene. Understanding these limitations, and how the specifications have improved over iterations of VR hardware, are important for understanding both the limitations of VR for use in vision research, and how vision research can help to specify the tolerance limits for when they will negligibly affect the perceived reality of a display. Therefore, this section will discuss different display types with increasing levels of immersiveness.

1.4.1 Non-Immersive 2D Display Systems

A television, computer monitor, and even a mobile phone screen are all display systems which can present high resolution 2D images. Significant advances have been made in display technology in recent years, which have substantially improved the viewing experience in terms of realism: ultra-high-definition, motion interpolation, and high dynamic range are all features seen more and more frequently in the specifications of screens and monitors. With an enhanced viewing experience comes the increased ecological validity of any experiment's findings, however the limitation of two-dimensional stimuli limits the applications of these types of screens. Specifically, the experience of viewing an image on a screen, will be that of viewing a picture, and the viewer will not be 'fooled' into believing that it is a real scene (Rogers, 2011).

When measuring the impact pictorial cues have on perception, 2D displays are often an appropriate way to present the stimuli. Being unable to manipulate binocular cues means that testing perception of distances is near impossible, but testing shape perception is feasible. Even though 2D stimuli, such as paintings or images presented on a monitor, are physically flat, a 3D percept can be obtained by viewing through other depth cues than binocular disparity: this can be referred to as *pictorial relief* and was coined by sculptor Von Hildebrand (1905). Pictorial relief refers specifically to the impression of depth within the scene: the depth perceived within the scene is distinct from the physical reality of the screen, page, canvas.

Measuring perceived depth within 3D displays is intuitive, but measuring it in a 2D display is more challenging. Koenderink et al. (2001) proposed many ways that pictorial relief can be measured from 2D displays; Koenderink and Van Doorn (1995) tested the perceived depth of a 2D stimulus by overlaying a small target at specific locations on the stimulus. The target was a gauge figure, which consisted of a circular base and a stem protruding 90° from the centre. In each trial, the gauge figure would be fixed at a location, while participants made adjustments to the orientation and extent of rotation (slant and tilt), which portrayed the shape perceived at these points. This method is ideal for investigating perceived depths within 2D scenes, as it is versatile to many different types of stimuli. The types of cues available on these types of display do not include binocular cues, so the interpretation of the findings in such experiments (the way cues are combined to perceive depth) may not be representative of the way depth is perceived in 3D scenes.

1.4.2 Non-Immersive 3D Display Systems

Any 2D display can be a 3D display by using anaglyphic colour filters: this technique allows for a stereoscopic image to be perceived at the loss of colour. Polarised glasses restrict the light that can reach each eye independently, which also results in a 3D percept, but the picture quality is compromised due to one screen presenting both the images

which are being delivered to each eye simultaneously. Two images with binocular disparities, presented very rapidly in succession, or concurrently, are the principal requirement for many current 3D display systems. Adding the potential for binocular vision to a 2D stimulus through the use of shutter glasses and a 3D stereoscopic monitor is another way to elicit binocular viewing.

Manipulated versions of stimuli can be presented using stereoscopic displays in order to isolate the contribution specific cues have on observer's responses. An example of this is the random dot stereogram, in which a random configuration of dots are presented at set distances so that when combined, the low-level shape of a stimulus can be observed from binocular disparity alone, and is void of all other visual cues. This approach also allows us to ascertain the accuracy of perceived depth from different combinations of cues, and the stimuli requirements that determine the contributions of each cue.

In principle, a binocular display reproduces the binocular images that would be experienced when viewing a particular scene. For this to be achieved, the image capture or creation, display, and observer positioning require accurate calibration and measurement. The perceived size and shape of a stimulus will be dependent on the physical distance of the stereoscopic display and the observed properties of the stimulus can be incorrect when the display is positioned at the wrong physical distance (Masaoka et al., 2006; Woods et al., 1993). This can result from a mismatch between the actual and intended retinal image, if calibration and positioning are not controlled. Often a chin-rest will be used to hold a participant's view constant and at the correct location relative to the screen (Asher et al., 2019), although ideally a bite-bar would ensure accurate and stable positioning (Backus et al., 1999; Koenderink, 1999; Paige et al., 1998). While this can be restricting and even invasive for participants, it does ensure the stimuli are presented as intended by the researcher. The use of these accessories allows for accurate positioning on the retina to which it is projected, as well as accurate luminance properties. While these features are vital in some experiments, the lack of natural viewing and behavioural constraints are limiting for others.

Psychophysical experiments have been coded in programmes such as MATLAB, utilising the capabilities of stereoscopic monitors and shutter glasses in order to present 3D stimuli (Georgeson et al., 2009; Hibbard, 2008a; Hornsey et al., 2015; Scarfe & Hibbard, 2013). With the use of a 3D monitor and its associated shutter glasses, it is possible to programme an experiment with basic 3D stimuli with tasks such as a two-alternative forced choice set-up. However, creating these types of experiments requires coding knowledge in order to write scripts in the software.

1.4.3 Immersive Environments

The *Ultimate Display* has been described as a virtual room where its sensed contents are controlled by a computer, but the virtual stimuli within are perceptually real (Sutherland, 1965). The realness of the stimuli are believable in terms of the sensory responses elicited, including visible consistencies and haptic feedback. Sutherland compared this room to the Wonderland into which Alice walked.

In companion, Turing (1950) stated that at some point in the future, computers will compete with humans in all intellectual fields. This was stated with regard to artificial intelligence, but Banks et al. (2016) put forward the idea of this concept in terms of displays: artificial stimuli competing with physical stimuli, so much so that a viewer cannot tell if an environment being observed is real or virtual. Zhong et al. (2021) built a multi-focal display system in order to test whether this was possible, given the capabilities of 3D rendering currently available. A multi-focal display system is one where there is more than one location in egocentric distance from the observer, that the observer's lens is possible to focus on due to presentation of stimuli across these additional focal planes. In the study for which this set-up was built, participants were tasked with identifying when a presented scene was artificial or physical; the multi-focal system was found to be identified as being different from the physical scene in 44 % of trials. Almost half of the time were the two displays determined to be the same: this set-up identified how advanced display technology is, and identified that cues such as focal blur may in

fact be more important than accommodation in perceptual realism. Multi-focal devices were created before this one (Hoffman et al., 2008), but this experiment best presents the importance of specific cue and cue-conflicts.

A device which allows the user to believe they are in the virtual scene presented, rather than a passive viewer, is the direction that VR is heading, with the aim of producing a truly believable albeit simulated experience. The sensation of dissociating from the physical world and being fully engaged in an alternate one (physical, virtual, or imagined), has been referred to as *immersion* (Agrawal et al., 2020). Very similarly, and often interchangeably, *presence* refers to the believability that the virtual stimuli presented are experienced as being real (Lee, 2004). The subjective tropes of realism and immersion are arguably enhanced in VR relative to other displays, as there are features of these devices which allow for a subjectively more engaging experience. Good VR should include the ability to interact with the virtual environment, which would be useful for an unlimited amount of research applications.

VR devices are commonly portable HMDs which present slightly different images to each eye for a binocular viewing experience. Alternatively, the Cave Automatic Virtual Environment (CAVE) is another a VR set-up where, instead of a headset, the virtual environment is projected onto walls surrounding the user and, through the use of 3D shutter glasses, simulated depth and distance can be perceived. Both HMDs and CAVEs have both recreational and specialised uses therefore advancements in these types of technology are both important.

1.5 Development of Virtual Reality

The advancements in VR technology over the past decade have been extensive, and there is no doubt that this trend of improvement will continue for the foreseeable future. Established companies such as Microsoft, HTC, and Sony have invested into the production of HMDs which have captured the attention of many disciplines, despite the idea

of wearing technology being around for decades. These however, have only recently become commercially available for the consumer market; prior to this, headsets were very expensive so were mostly used in laboratories and not something a household would typically invest in. VR can be an effective means of studying vision, as it is possible to program paradigms which have been validated in physical space, with the added control over extraneous variables. For example, sectioning tasks have been used to test distance perception in outdoor environments (Purdy & Gibson, 1955; Rieser et al., 1990) which can now be tested in VR whilst being able to manipulate specific factors such as lighting with ease.

1.5.1 Requirements for Effective Virtual Reality

As the use of 3D displays has widened, the benefits and problems associated with them have become clearer (Held & Banks, 2008). In order to enhance the effectiveness of the applied functions of VR, the basic components necessary in establishing a naturalistic viewing experience need to first be nailed down. VR promotes a more naturalistic viewing experience than 2D displays, as more parameters are able to be coded into experiments to be tested and/or held constant. Because of this, results obtained in the naturalistic virtual environments of VR would be predicted to be more generalisable to real-world viewing, than results from more restrictive 3D set-ups.

As previously described, accommodation is a biological cue to distance perception in which the lenses of the eyes adjust to allow for objects at any one distance to be in full focus at a given time. In VR headsets, the screens are positioned very close to an observer's eyes, and as the form factor of HMDs becomes more efficient, the size will further decrease. In commercial headsets there are lenses positioned between the eyes and the screens which extend the focal distance to a distance beyond the physical screens, which is important due to the human eye struggling to comfortably accommodate at close distances (Shibata et al., 2011b; Yang & Sheedy, 2011), and the minimum distance

of accommodation has been shown to be positively correlated with age (Duane, 1922). The use of such lenses allows the focal distance to be increased, which is beneficial for the visual system, however a singular lens in front of each screen results in a singular distance for all stimuli to be positioned at. In natural viewing, the accommodative response aligns with the presentation distance of observable objects, but in consumer VR devices the singular focal distance often results in a conflict between the biological cues of accommodation and vergence. Purpose-built research apparatus have been developed to allow for stereoscopic displays with multiple focal distances (Aghasi et al., 2021; Hu & Hua, 2014; Liu & Hua, 2010; Mercier et al., 2017; Watt et al., 2012) which is desirable for accurate and comfortable viewing, but is currently unattainable in portable, consumer HMDs. Advancements in this domain, for example lenses allowing for multiple focal distances, would improve the subjective experience of the observer by reducing the fatigue arising from the accommodation conflicts (Hoffman et al., 2008), and provide enhanced ordinal information over the depths within the virtual environments (Mon-Williams & Tresilian, 2000).

The graphics and rendering are also key components of any visual display. The size of a device limits the capabilities of the internal screens: many are not able to produce the dynamic range and gamut of colours which are possible in non-immersive displays. ITU-R Recommendation BT.2100 defines various aspects of high-dynamic-range screens, which represents the location of the primary colours in colour space; the coordinates of the colours for this version are more encompassing than previous versions, which were used in lower specification displays and are still used in many of the current VR headsets. This is an area which is lagging behind previous technology and so VR falls short for these specifications.

The field of view (FoV) is also limited by the size of the headset, however the documented increase in this from newer headset releases (see *Section 1.5.3*) is edging toward a degree which is comparable to natural viewing. The horizontal FoV in natural viewing for both eyes is around 210° with a binocular overlap of around 120° (Veron et al.,

1994). The documented increase in FoV will undoubtedly result in a preferred viewing experience compared to more restricted viewing, as FoV modifications has been shown to impact this (Arthur, 2000; Fernandes & Feiner, 2016). Larger FoV would require a larger device and therefore would lend itself to being also heavier and bulkier, which is not favourable for a consumer device.

The compression of space, which has been previously discussed, has been explored in terms of the capabilities of the screens within VR devices. The quality of the graphics (Thompson et al., 2004), FoV (Creem-Regehr et al., 2005; Knapp & Loomis, 2004), and stereo information (Willemsen et al., 2008) have been suggested to be potential causes of such a bias in distance estimates. A definite conclusion is yet to be drawn, however the rendering of the virtual scenes and the graphical abilities are areas of frequent exploration and improvement, and as such are requirements for effective VR.

The addition of binocular information contributes to the quality of perceived metric depth (Hibbard et al., 2017a). HMDs generally have a separate screen region visible to each eye, and lenses positioned in front of the eyes. The short distance between the eyes and the screen, of just a few millimetres, limits the resolution that is possible, so that individual pixels are often visible to the observer; this pixelisation is not conducive to a believable alternate environment. The image quality for 4k TVs have a recommended viewing distance so as not to see pixels: the pixels, when viewing ultra-high-definition TVs at the correct viewing distance, are below 1 arc minute. HMDs will eventually have a pixel density high enough that observers with perfect vision will not be able to detect any individual pixels, however at the moment the approximate pixel density of consumer headsets is well below this, and can be calculated:

$$\texttt{PixelDensity} = \frac{\texttt{HorizontalPixelsPerEye}}{\texttt{MonocularFoV}} \tag{1.6}$$

Using *Equation 1.6*, the density of the consumer headsets on the market at the time of writing this are all well below what the human retina can identify, which is 60 pixels per degree. Carefully calibrated equipment which is used in research facilities, such as

the 8k Dimenco Simulated Reality screen (Dimenco, 2022), have a much higher resolution than consumer devices, however improvements in the market are increasing the functionality of readily available devices.

The ability to interact with a virtual environment presents an issue in itself, with the conflict between the virtual and physical worlds. Although presentation of a stimulus may be perfectly achievable, when observer motion is incorporated, what happens when the observer moves into the space being occupied by the virtual object? In an ideal set-up there would be either haptic feedback or the stimulus would incorporate some laws of physics by moving when contact is made. This is still an issue with current models of VR, although software developments and equipment add-ons are increasing in their appeal to consumers: HaptX have created gloves which give sensory feedback when used in combination with VR (HaptiX, 2021). Each glove has over 100 points of feedback and can deliver up to 22 newtons of resistive force per finger, meaning that these would be ideal for training surgeons or creating more of a wow-factor during game play. These are not priced for use by a general consumer, but for specialised use.

Motion tracking is another feature required for interaction, where the observer's head position is taken into account so that the correct ambient optic array of an environment can be simulated. Any movements made by the observer will result in an altered viewpoint and therefore the environment will be rendered in accordance with the adjusted viewing position. The lack of updating to take account of the users position in other displays is one way other displays elicit a lower sense of immersion, compared to VR devices where this is used and found to enhance the experience (Angarita et al., 2019). Accurate tracking is necessary for this: a tracking system which inaccurately tracks would be a redundant component of the device. Motion tracking on controllers, especially ones such as the HaptX, are necessary for an effective VR experience.

To use VR to its full capacity, a larger area of space should be made available than what has previously been required for first-wave VR HMDs. A limitation of non-

immersive stereoscopic displays is that they are unable to measure exploration and behavioural components: although a selling point for HMDs is that the user can move around while using the device, there is still the limitation of the wires connecting the headset to the PC or console, meaning there is a potential tripping hazard while absorbed in the virtual world. Of course the length of wires and size of physical environment depends on the use of the equipment, but for some experimental purposes it is necessary to walk in the environment. Exploration, or blind walking for example would require the equivalent of a corridor in real space; to use this as a method for studying the effects of specific variables on distance perception would be difficult when the wires may not reach the entire length of the distance, and the area of the room would also be a constraint. The use of a VR device attached to a laptop and carried by a participant is a laborious workaround, but a tracking system which covers the whole extent of the area accurately would be favourable. To safely experience VR this is something which needs to be conquered: the Samsung Gear has overcome this problem, by using a mobile phone as the insertable console so there is no need for power cables or connection wires (Samsung, 2015). Oculus took a step in this direction in the creation of the Oculus Go in 2018, another all-in-one VR headset, but this lacked in the processing power and tracking capabilities of the cabled devices. Since then, the Oculus Quest has taken its place as a wireless alternative with increased processing capabilities (Oculus, 2012). In order to create a fully immersive virtual environment, one should be able to behave as they would in a natural environment, and not be limited to a small movement perimeter.

The need to calibrate the equipment spatially, in terms of observer metrics and internal screen and lens location, are also essential. Observer metrics such as visual height can directly impact the perception of an environment: a study by Dixon et al. (2000) found that viewing stimuli in an immersive environment at differing heights affected judgements of stimulus size; the heights tested included mismatches between eye-height assumed by the system for rendering, and the eye-height of each individual participant. Similar effects of height have been found on distance estimates (Corujeira

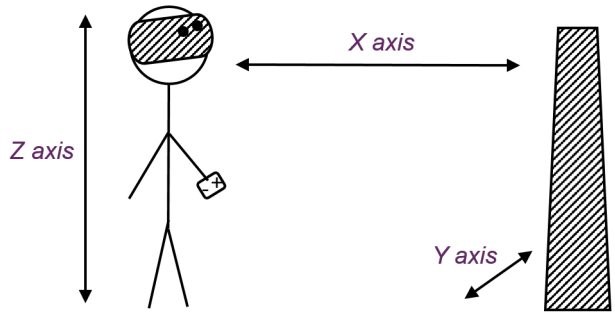


Figure 1.9: Depiction of the axes used throughout this thesis. The X axis within the headset depicts the saggital plane, egocentric distance from the observer; Y axis represents the transverse plane, left to right; Z axis represents the coronal plane, up and down.

& Oakley, 2013; Cutting & Vishton, 1995; Leyrer et al., 2011) and immersion (Deng & Interrante, 2019; Karhunen, 2020; Keskinen et al., 2019). *Figure 1.9* depicts the axes used throughout this thesis, any increase or decrease in height of the observer along the Z axis would result in a different viewing angle to the stimulus. Although adjustments in one's height changes throughout the day, from laying down to standing up to being seated, the undesired errors arising from observing a virtual environment which is calibrated incorrectly along the Z axis from a standing position can easily be eliminated through the ability to set one's height or even a tracking system on the device.

Another individual difference which directly affects the use of stereoscopic displays is the IPD. The population average IPD of 63 mm (Dodgson, 2004) can be used in experiments where the effects of this distance are not pertinent to the research. Differences in IPD do have an effect on the amount of depth observed: Hornsey et al. (2016) systematically varied the effective IPD of participants viewing a 3D scene through the use of computer rendering. The set-up allowed for reductions and increases of participants' IPD to be simulated: the distance between the two vantage points presented to each eye was scaled by 0.5, 1, 1.5 or 2 times their actual distance, and observed through shutter glasses. It was found that by increasing the effective IPD, the depth perceived by participants also increased. With this in mind, some VR headsets allow for the distance between the lenses/screen in the HMD to be manually altered to match the user's IPD, however whether or not casual users actually understand how important this is for their

viewing experience is questionable. A potential way to overcome this issue would be for manufacturers (or experimenters) to set the distance in their headsets to be the population average, so as to cater for majority of individuals (68 % of the population within one standard deviation). This obstacle could explain why some individuals encounter a different perceptual experience in VR; for example in *Figure 1.2*, if an observer's natural IPD matches IPD_1 but used a headset with a larger distance of IPD_2, the depth of objects would be larger than what they would be used to. Visual fatigue and motion sickness may also arise from this mismatch, and act as a deterrent from future use of the technology (Hibbard et al., 2020), which is why an effective headset should include some sort of manual adjustment to account for those with IPDs outside of one standard deviation of the population mean. Differences in the IPD between male and female observers has been shown to vary (Dodgson, 2004), so for a device to be used appropriately by the majority of the population it is essential that this setting be able to be modified.

1.5.2 First Attempts of Virtual Reality

First-wave VR systems included the likes of the Sword of Damocles by Sutherland (1968), which was the first attempt at a VR headset. The stereoscopic display was able to present virtual, wire-frame rooms and used an ultrasonic head position sensor to track head movements via a ceiling arm. This was the first step towards the creation of the *Ultimate Display* (Sutherland, 1965) which current developers are still trying to achieve, where the combinations of sensory feedback and computer rendering perceptually place an individual in a new environment.

One of the first consumer-level HMDs was the Forte VFX1 produced by Forte Technologies in 1995 (MINDFLUX, 2006). With an introductory price of $695 and running requirements (a Windows 95 computer with 20 kb conventional memory) it would appear that the target audience would be the general public, however when taking inflation into consideration, the 2021 price would be approximately $1,261. Forte made this com-

patible with a total of 100 games and applications; the technical specifications of the device included built-in speakers and microphone, stereoscopic Smart Visor with 263 x 230 pixel resolution, capable of producing a 35.5 x 26.4° (45 diagonal) FoV, and a gaming controller. Remarkably, the IPD was adjustable on this early device and had a three degrees of freedom from the sourceless tracker (Neale, 1998). Although this product was aimed at the gaming industry it was adopted by fields more able to deal with the high price. For example, Alvarado et al. (1999) used the device for architectural purposes in order to get a virtual representation of how different timber structure designs may appear, prior to physically building them. Although their study was not designed to assess the quality of perceptual experience in the device, the findings showed a 20 % higher rating in the functionality, stability, and aesthetic of the virtual presentation of their schemes when compared to traditional method of sketches, indicating a preference for VR even at this low level of presentation. This product was not available for long , which may have been as a result of the lack of compatibility with the higher quality 16-bit games and 3dfx graphics cards that were appearing around the same time as the VFX1.

Other, less discussed attempts in the second wave of VR headsets, include the VictorMaxx Cybermaxx which was also a 1990's development with 505 x 230 resolution and 56 x 42° FoV (Healey, 2016). With a screen resolution of almost double that of the VFX1, and costing a similar amount, rapid developments in the technology suggest that perhaps the reason for VR not becoming commonplace in this era was societal as opposed to quality being the limiting factor of uptake. On the other hand, the idea of immersive environments were introduced to the general public and making appearances in the form of the Recreation Room from the television franchise Star Trek, which made its first appearance in 1974 and has continued to be a key feature in many of the individual series such as The Next Generation (Roddenberry, 1987) and Voyager (Berman et al., 1995), while being referred to as the *Holodeck*. In addition to this, the 1990's produced many VR-based films including The Lawnmower Man (Leonard, 1992), eX-

istenZ (Cronenberg, 1999), and the cult classic: The Matrix (Wachowski & Wachowski, 1999). Perhaps if the efforts has been invested into the development of CAVE's instead of HMDs there would have been more success in this era due to this media influence, however VictorMaxx and its products came to an end when it was sold in the same decade.

1.5.3 Current Virtual Reality Models

The specifications required for a computer to run some of the consumer HMDs vary between devices, but to be able to run more than one brand of device for any given experience, a PC with powerful graphics hardware is needed. Although this would unlikely be an issue for research institutions, many casual users might not have the funds for this equipment and therefore may not be able to use them to their full potential and replicate the findings of studies with such high-end computers. It is important for casual users to take this into consideration, but also for researchers to understand the consequences of lower-end hardware for the user's experience. To circumvent this obstacle, some brands are moving toward untethered devices. During the third wave of VR, here are the specifications of some of the commonly used devices at present.

The Oculus Rift is a prominent consumer HMD which was first released as a Developer Kit (DK1) in 2013 and, as the name suggests, was aimed solely at developers rather than consumers. This had a screen resolution of 640 x 800 pixels in each eye, refresh rate of 60 Hz and 110° FoV. Although also having three degrees of freedom tracking, the DK1 used a 1000 Hz Adjacent Reality Tracker which allowed for lower latency rates than a HMD from the first wave of VR: the VFX1 (Brown, 2021). The second version of this device (DK2) was developed, followed by a consumer version (CV1) which was released in 2016 with further improved specifications: 1080 x 1200 screen resolution in each eye, a 90 Hz refresh rate, 5 x 11 feet (1.5 x 3.4 metre) tracking area, and six degrees of freedom positional tracking system. This device requires a PC with minimum 8 GB RAM available (Oculus, 2012). In this headset, the IPD can be adjusted through

a manual slider at the bottom of the headset, with a range of 58 to 72 mm (Knight, 2019).

HTC's consumer version of the Vive was made available in 2016 and, on the surface, had the same technical specifications to the Rift with regard to the resolution, refresh rate, and FoV (Vive, 2019). However, needing a lower capacity of PC RAM (4 GB) and having a larger tracking area (15 x 15 feet), the HTC was therefore a main competitor for Oculus. An additional user experience feature of the Vive was the Chaperone system which, when manually set-up, would provide a warning the user about the physical boundaries of their physical environment. The Vive was powered by SteamVR, meaning that the thousands of VR games created by third parties were easily accessible and allowed for updates. The IPD on this headset is able to be adjusted manually, for each individual observer (Vive, 2021a).

Samsung created a HMD which utilised the technology within the user's phone, instead of having a built-in screen, speakers, or location sensors. The Samsung Gear VR was released to consumers in 2015 and had a 101° FoV, however it did require a Samsung phone to be inserted to work (Samsung, 2015). Due to this, the resolution of the display was dependent on phone specification; additionally the headset was only compatible with the latest Samsung models, which was a limiting factor for potential users. The screen resolution of compatible devices were all high-definition, with recent models exceeding this at 2400 x 1080 pixel resolution (Samsung, 2021a, 2021b): far surpassing the display capabilities of the Rift or Vive. Unlike these headsets, the distance between the screens (simulating the distance between the eyes) of the Gear was fixed at 62 mm due to the phone set-up, and Samsung were reliant on external content creators for software; most content which can be downloaded for use on the Gear was supplied by Oculus Home. Given that the inserted phone provided all the tracking and movement sensors, as well as programme information, this wireless HMD opened the scope for potential uses of the Gear which were unobtainable from using a wired headset.

Nowadays, multiple systems are available to the general public, but this thesis uses

two of the prominent devices at the time: the HTC Vive and Oculus Rift. The specification details of these devices are outlines in *Chapter 3, Table 3.1*. There is a lack in research comparing the current VR headsets available to both consumers and the research industry, potentially because there is no need to compare the newer models to their predecessors due to the implicit assumption that the technology is only going to improve. However there is a need to compare the HMDs which are of the same generation but produced by different companies, such as Oculus and HTC, because the individual research and development teams working on these devices may place emphasis on different factors. This could show significant differences on the subjective experiences of users. One chapter of this thesis will directly compare two headsets in their presentation capabilities.

1.5.4 Physical Space and Virtual Reality

A body of research which is essential to the development of VR technology is that which compares performance in both virtual and physical environments. A series of studies will be explored here which have tested this.

Hartle and Wilcox (2022) conducted an experiment which aimed to determine how viewing distance and cue conflicts impact depth perception. A depth estimation task was created, where textured half-cylinders were presented to participants in both a physical testing space, and an identical scene presented virtually. To address the issue of cue conflicts, the virtual scene was presented in both a VR headset (in which conflicts such as accommodation-vergence are present, as explained in *Section 1.3.1*) and a mirror stereoscope (conflicts were lessened), as well as the physical scene with 3D printed stimuli which replicated the virtual stimuli in terms of size, contrast, and texture but with zero visual conflicts. Participants indicated the depth of the cylinders in each trial by adjusting the aperture between thumb and forefinger on a touch-sensitive strip. Participants took part in all viewing parameters, where viewing device (or lack thereof) was counterbalanced, and two presentation distances of the cylinders were experienced. Performance

in the task was similar in the two virtual environments, for which the authors concluded that the optics in the display units had little impact on perceived depth. However comparing performance to that in the physical environment, accuracy was much lower when the stimuli were virtual. Given the lack of significant differences of participant responses in the two virtual environments, specific cues which were not able to be controlled for, such as retinal blur, may be the reason accuracy was lower than veridical.

The use of cues in natural viewing may not correspond with the way these cues are utilised in unnatural and/or virtual viewing. A three-part experiment by Frisby et al. (1995) investigated the impact of visual blur in depth perception. The paradigm involved manipulating the cue of blur that was able to be utilised by participants, when making estimates about the depth within presented stimuli. Results indicated that blur was indeed an important cue for depth, and performance was different between the different sets of stimuli. From this, the authors stated that caution needs to be taken when drawing conclusions in one testing set-up, and applying these conclusions to another set-up. Namely here that results from observed depth of physical stimuli are not necessarily the same for observable depth in stereograms.

A review by Loomis, Knapp et al. (2003) builds on this. Here, the limitations of many traditional methods used to calculate egocentric distance perception are discussed, with regard to both physical and virtual environments. The author notes the variability in estimates that may be a direct result of the methods used, as well as the specific stimuli and equipment.

An influential study by Kunnapas (1968) aimed to investigate the effects of adding more cues into an observable scene. From a reduced cue environment, the author gradually added available cues until the fifth environment presented them all without manipulation. This experiment was conducted in a physical environment, with verbal estimates from participants about the distance to a stimulus in each trial. Distance of the stimulus was manipulated between trials, and the results showed that estimates were not accur-

ate, with performance decreasing as presentation distance increased. Estimated distance to the stimulus was consistently reported as being shorter than the actual presentation distance in the reduced cue conditions, but switched to in the full cue environment with estimates being further than accurate. A similar technique is used in one chapter of this thesis, where a series of environments incrementally add to the observable cues, however in this thesis, this has been conducted in VR.

Similarly, Leyrer et al. (2015) conducted a height estimation task in VR, comparing the results from a reduced cue environment with a rich cue environment. The height of the observer was systematically adjusted, along with the environment, as participants were tasked with estimating their own height in the virtual environment. In the reduced cue environment the available cues were stereo, texture gradient, and linear perspective; scene clutter was added to incorporate familiar size in the rich cue environment. Adding cues in this experiment was found to reduce variability in estimates, but only marginally increase accuracy. Evidence suggests cues are combined similarly in physical space and virtual space, as results in similar tasks are alike.

For another direct comparison, von Castell et al. (2021) conducted a distance estimation task while varying the eye-height of observers in both VR and physical space. Participants either sat down or stood up while making estimations on the height, width, and depth of an observable environment, which allowed for postural cues to also be controlled for. Distance judgements were found to be unaffected by postural cues, which meant perception remained consistent regardless of participants' posture. However when viewing height was adjusted separate from posture, as if the environment was being viewed from a seated position but participants were in fact standing, judgements on the dimensions of the environment were larger than when the viewing height was corresponding to being stood upright. This task was replicated in a virtual environment and a physical space, and yielded results which were comparable. The authors concluded that eye-height information is processed similarly in both physical and virtual environments, which is why height of an observer in VR should be congruent with their actual height,

regardless of their posture, if accurate performance is desirable. This was taken into consideration throughout the experiments in this thesis.

1.6 Inconspicuous Applications of Virtual Reality

Because it is such a versatile technology there are many potential uses for VR, reinforcing just how important it is that research into it be as comprehensive as possible. One prominent ways VR is utilised is for entertainment, although its uptake is by no means restricted to this field: the specifics for each section can be applied to different industries, as the same ideology for an immersive experience in an alternate environment is striven for. Discussed below are some less common, yet momentous applications of the equipment. The quality of visual presentation within the headset, as well as form-factor, are of high importance due to the nature of its applications.

Depth and distance perception are particularly important in certain applications, more so than the mere aesthetics of the virtual environments. An example of two applications would be for recreational and training purposes: 2D gaming is more than capable at presenting desired depth qualities through pictorial cues alone, so the addition of binocular functioning within a VR headset is less critical than needing disparities and other binocular cues when doctors use the same devices in training for surgical procedures (McKenna et al., 2008; Shimotsu & Cao, 2007). Discussed here are other instances where the equipment is currently used, the necessity of accurate perception in each of these applications is generally subjective, with everyone using the equipment likely wanting to have a perceptually accurate experience. However, there are some instances where it is categorically more important to have a faultless perception.

1.6.1 Distraction from Pain

Distraction from pain, in many medical procedural cases, is vital for a successful outcome. Pain requires attention (Eccleston & Crombez, 1999) and inherently attains attention, therefore any attention interference should reduce pain felt. Vasquez et al.

(2017) compared subjective ratings of pain and physiological measurements of stress in 44 women during gynaecological procedures, where one group experienced a VR distraction and a control group did not. The VR distraction was a clinically validated environment in which the observer could navigate a virtual environment through the use of a joystick. Both groups had their pain and physiological measurements taken at three points of the procedure: before, during, and fifteen minutes after. The self-report scale revealed a 24 % decrease in pain from before to during for the VR group, compared to a 13 % increase in pain experienced by the control group. A significant reduction in pain was found after the final report of pain in both groups. The physiological data set revealed that the VR group had lower heart and respiratory rates than the non-VR group before procedure, expectedly rates rose during, and finally the VR group showed a more drastic decline in these rates than the control group did fifteen minutes after. The key conclusions of this study was that not only can VR distractions ease procedures, but it can also reduce anxiety experienced by the patient before hand.

A similar study by Maani et al. (2011) used a self-reported rating scale for the amount of time patients having debridement performed on them (removal of damaged tissue or foreign objects from a wound) thought about the pain they were experiencing. A subsection of patients who rated the worst pain felt in the experience (seven out of ten or higher) were compared from both the VR distracted group and a non-VR group: it was found that patients in the VR condition experienced significantly less time thinking about the pain. A feature of VR is to present an immersive, alternate environment, and by doing this it needs to be as captivating as possible. Because of this, VR experiences can be used as successful distractions from pain or discomfort in clinical settings.

1.6.2 Educational Training

Seymour et al. (2002) explored VR as source of training new motor skills by comparing two groups of surgeons: the first were trained for a laparoscopic cholecystectomy (removal of the gallbladder) via the standard procedure, and the second were trained

using VR. The standard training technique for surgeons learning this procedure is usually either practising on pigs or cadavers (Hunter et al., 1994; Nagase et al., 2020). When these surgeons then performed the surgery on a patient, two blind investigators scored them: the non-VR surgeons were five times more likely to injure the gallbladder or surrounding tissue, and the VR surgeons were 29 % faster than the other group. The results showed that operation techniques were better learned with the use of VR and the conclusions of the study were that this type of training could be implemented for more sophisticated types of surgery.

VR can be used not only to teach clinically important skills, but social ones, also: a review by Parsons and Mitchell (2002) looked at how VR has been used to teach people with autism and found that it allowed for an environment where the participants could learn and practise skills. The conclusions made were that VR is very appropriate for this in that it allows for a safe environment for rule learning and repetition of tasks, as opposed to techniques which are less immersive such as learning with a computer or tablet.

In addition to learning new skills, VR has been used to train fire fighters in safe space where incidents are able to be repeated for practice in what would otherwise be high-risk (yet low incidence) situations. Narciso et al. (2020) measured perceived presence, heart rate, stress, and fatigue in both a virtual and physical environments where fire fighters were tasked with controlling a fire. This experiment found that the set-up was able to successfully display the scenario, in that participants did not experience any motion-sickness yet were able to experience high levels of presence in the virtual environment. Similar software programmes have been created to train the emergency services in similar scenarios, including Bodyswaps (2021), FLAIM (2021), UbiSim (2021).

In order to learn or refine skills, a life-like environment is essential. Learning in a controlled environment such as this avoids the negative ramifications of any incorrect decisions and is valuable to learners of all skill levels.

1.7 Attitudes Towards New Technology

Even if all technical components of a new display system are achieved by manufacturers, this does not mean that the equipment will be used in research, let alone accepted by the general public. The reasoning behind individual's decisions needs to be understood, because steps can be taken to encourage people to want to use it.

The Technology Acceptance Model (TAM) of Davis et al. (1989a) aimed to explain consumer behaviour and the likelihood of individuals welcoming the use of a specific example of technology. This model suggests that by assessing one's perceptions of how useful and easy to use the technology in question is, it is possible to gauge their attitude towards using it, as well as their behavioural intention to use it. In turn, these attitudes and intentions are expected to determine the likelihood of the individual actually using the technology in the future.

The Virtual Reality Hardware Acceptance Model (VR-HAM) (Manis & Choi, 2019) was built as an extension to the TAM, specifically for the acceptance of VR devices. Manis and Choi hypothesised that there were additional factors contributing to the use of a device which were not covered by the TAM, especially as there were advancements in technology that were not considered when the TAM was created. Due to a major current use of VR being gaming, the predictor of perceived enjoyment was added; as were curiosity, price, and prior experience with the device. The ability to predict the popularity of a device is important for manufacturers, but research into this also provides an insight into the decision making process on individual and societal levels.

A topical consideration surrounding the adoption of VR is the potential for this technology to collect and exploit personal data. The European Union's General Data Protection Regulation (GDPR) outlines how organisations must collect and protect their customers' personal data, with financial penalties for any organisation who does not comply (European Parliament And Council Of The European Union, 2016). Although

this has been implemented, it has been suggested that most organisations in the USA and Canada are not compliant with the GDPR (Miglicco, 2018), where the regulations are not in operation. VR devices have the ability to collect much more personal data than other online gaming devices; for example the distance between an individual's eyes and the height of the person using the headset. Not limited to the personal metrics, but information on the way individuals explore and interact with environments may also be collected. Although there is a logical reason for this data to be collected, so that a comfortable viewing experience can be provided, there are questions surrounding the data afterwards. What happens to that data after, who has access to it, and how long is it stored for are all points for concern when the organisations who build the equipment may not be compliant with GDPR.

Concerns are not unjustified when it comes to the data collected through VR devices, given that many of the hardware creators are data-driven in nature (Henriksson, 2018). Olivi (2019) outlines the main privacy issues related to VR use, which include the processing of biometric data, informed consent, and children's data if minors use the devices. These metrics, coupled with the customisable avatars and behavioural data obtained via gameplay, no doubt have an impact on the attitudes of using these devices.

From this research, the link between individuals' attitudes towards VR and their intent to not only use it in the future, but also purchase the device themselves is apparent. Attitudes are not fixed and are able to be changed (Petty et al., 1997), especially when the price, usefulness, and risks of using a device are ever changing. Because of this, it is important to frequently assess the attitudes towards the new hardware developments to not only provide more detailed insight into the theory behind predictive modelling, but also to inform the manufacturers of what consumers deem important. User research enables understanding of barriers which prevent consumers from using a device, and this informed approach can be used to overcome these obstacles. These topics are investigated in one study within this thesis.

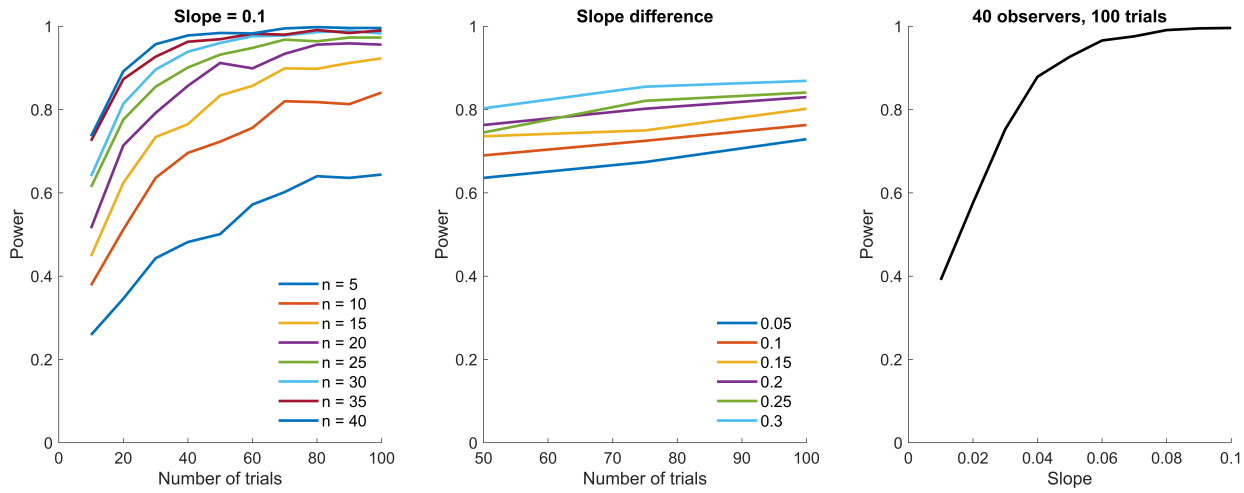


Figure 1.10: Power calculations for the experimental chapters in this thesis, where the number of participants and number of trials were manipulated in a simulation.

1.8 Presented Research

1.8.1 Power Calculations

The studies presented in the experimental chapters of this thesis were designed to quantify the perceived distance in VR devices. This was evidenced by the slopes of specific regressions conducted on performance in a number of tasks. It is possible to quantify the power of our methods post hoc, to detect when slopes differ from zero, and differences in slopes between conditions.

Based on pilot data, the value of the correlation between rendered and perceived distance was approximately equal to the slope of this function. Additionally, the standard deviation of the slope across participants was also equal to the mean slope. These relationships were used to determine statistical power.

In the first simulation, the power necessary to detect a slope of 0.1 or higher was calculated. The number of simulated participants was varied between 5 and 40, and the simulated number of trials per participant was also varied between 10 and 100. For each combination of these values, 1000 runs of the experiment was simulated. For each run, the slope of the relationship between rendered and perceived distance was drawn, for

each observer, from a Gaussian distribution which had a mean and standard deviation of 0.1. Rendered and perceived distances were then generated using a multivariate random number generator, with the slope specified for that participant, and the correlation equal to the slope.

A linear mixed effects model was then applied to the data for all participants in each experimental run, and the proportion of experiments on which a significant slope was detected was counted. This power measure is plotted in *Figure 1.10*. This shows high power to detect even modest slopes of 0.1 using the number of participants and trials in our experiments.

Following this, a similar approach was used to detect a difference in slope between conditions. In this case, a slope of 0.3 was modelled in one condition, and a slope in another condition that was larger than this by either 0.05, 0.1, 0.15, 0.2, 0.25, or 0.3. Data were simulated for 1000 experiments as above, but with these two conditions. An experiment with 40 participants was simulated, with between 50 and 100 trials per participant per condition.

A linear mixed effects model was fit to the simulated data, with distance as a linear predictor, condition as a categorical predictor, and their interaction included in the model. The power to detect a significant interaction term was calculated, indicating a difference in slope across conditions. This shows a power of greater than 0.7 to detect all but the smallest difference in slope.

1.8.2 Summary of Chapters

A selection of questions have been raised by the discussed research into visual perception, resulting in an abundance of potential research using VR equipment. Not only would it be useful to replicate published studies conducted on different types of 3D apparatus, but also adapt certain methodologies to explore other possibilities for specific results. Increased numbers of participants are used throughout this thesis compared to many

of the influential studies on which these experiments are based. The current thesis will explore five main topics, sectioned into one methods chapter and four experimental chapters. *Chapter 2* focus on the use of a games engine (Unreal Engine 4) for scientific experiment creation. This chapter includes a tutorial in which a basic VR experiment can be created. *Chapter 3* is a comparison of two current VR HMDs: the Oculus Rift and HTC Vive. A distance perception task was used to detect the hypothesised mis-perception of distances presented in virtual environments. *Chapter 4* explored the weighting of specific visual cues which are utilised during perception of distances through three experiments, including direct and indirect distance perception paradigms. *Chapter 5* investigated the influence directional lighting techniques have on distance perception, where observable shadows from target and clutter stimuli varied. *Chapter 6* took a more exploratory approach, whereby user's views on VR technology were probed.

Exploring the parameters which influence our interaction and perception of 3D environments is the basis for this thesis. The methodology used in the experimental chapters stem from validated techniques used in vision science: shape and size constancy tasks, a distance sectioning task, and an equi-distant positioning task, all programmed using Unreal Engine, and presented in VR headsets. The methods involve both direct and indirect methods of examining the perception of distance, and the games engine facilitated the isolation of the specific visual cues tested in each experiment.

The study which explored subjective perceptions was conducted via an online questionnaire, with closed and open questions to gain an insight into the viewpoints of participants. The questionnaire allowed for these subjective experiences to be explored, which is otherwise not possible from the quantitative, psychophysical experiments created for the other chapters of this thesis.

The four studies presented, and the tutorial chapter, also provide a basis for future experiments. Future research directions for which ideas are outlined in *Chapter 7*.

Chapter 2

Using Unreal Engine for Creating Quantitative Experiments

Primarily built as a development tool for games creation, Unreal Engine contains the essential and desirable aspects for building quantitative experiments with ease. Unreal Engine has been used to create the laboratory experiments presented throughout this thesis. This chapter provides a review of quantitative experiments and methods, justifies the use of Unreal Engine as an alternative to specialist programs, and outlines the set-up of many features which are used particularly in psychophysical and other quantitative experiments. In addition, a walk-through for creating a basic psychophysical experiment in virtual reality is given. Basic understanding of the way Unreal Engine operates is not necessary to have before using this chapter: key components and terminology are explained here to provide a starting point for any user aiming to use the engine in their research. For primary research see *Chapters 3, 4, 5,* and *6*.

2.1 Introduction

The aims of this chapter are to provide readers with a rudimentary understanding of Unreal Engine (UE), and to present a step-by-step tutorial to create a basic quantitative experiment which can be run using a virtual reality (VR) headset or computer screen. This chapter presents a background for UE, as well as a review of some of the techniques explained. This is not authored by a team of experts in UE, but by a psychology researcher using VR who has learned through doing and written this to share the tricks picked up along the way. It is thus written from the perspective of providing a gentle introduction to how a games engine such as UE can be used to create psychophysical studies to be run in VR.

The experiment to be created will be able to be viewed by participants in a VR headset. They will be positioned in a virtual room, the experimenter will instigate the experiment by pressing a button on the keyboard of the PC. A stimulus will then appear, the rotational direction of this will need to be memorised by the participant. The stimulus will then disappear, and reappear with a new, random rotation. Using the VR controller, the participant will be able to adjust the rotation of the stimulus until it matches that which was presented at the start of the experiment. A button press from the participant will save the adjusted rotation and automatically move onto the next trial. Five trial repetitions will occur, before the experiment ends and data is output into a comma separated variable file. This example thus covers some of the basic components required in an experiment, namely (1) a virtual environment; (2) ordering of trials and randomisation of stimulus parameters; (3) experimenter control of the experiment; (4) participant responses; and (5) saving the data.

2.1.1 General Experimental Set-up

The use of psychophysics in scientific research aims to provide reliable data through rigorous methodology. For example, psychophysics can be used to test the point at which a stimulus can be perceived, sometimes a sound (Fastl & Zwicker, 2006) or a

visual stimulus (Goutcher & Hibbard, 2014). A standard methodology for doing so is repeatedly presenting a stimulus with a variation of what is being tested, at small increments, while the observer responds within a two-alternative forced-choice procedure. The point at which the observer can identify the change in stimulus can then be measured and is referred to as the *discrimination threshold*, or *just noticeable difference* (JND). The point at which the observer perceives two variations of the stimuli to be identical on the dimension of interest, when there is equal probability that either of the two forced-choice options may be chosen, is referred to as the *point of subjective equality* (PSE).

In vision research, psychophysical experiments are frequently created through code languages such as MATLAB and Python, with linked equipment such as VIEWPixx (Hibbard et al., 2012; Meyerhoff & Papenmeier, 2020; Rohde et al., 2016; Sogo, 2017). Extensions used with this programming software often provide an active forum for users to share and discuss experimental design. Extensions such as Psychtoolbox and VideoToolbox for MATLAB, and PsychoPy for Python offer platform-independent control over timings, colour, and size scaling of stimuli when creating psychophysical experiments (Peirce, 2007; Pelli & Vision, 1997). These platforms greatly simplify the creation of experiments by allowing the researcher to focus on the design and creation of stimuli, rather than the details of the graphics hardware and operating system of the computer used to generate these stimuli.

These types of software allow the experimenter to define a flow of control so that blocks of trials can be ordered (or randomised), as well as one-off trials being able to be coded. These sorts of programs are ideal if one is competent in coding, or is able to collaborate with another who is competent; understanding the coding language is a limiting factor for these types of perceptual studies. Obtaining the apparatus and skills required to code in these computer languages certainly lessens the niche of potential users.

2.1.2 Virtual Reality in Vision Research

VR equipment is becoming more easily attainable by consumers and the hardware specifications are also increasing. Exploration into VR technology has dated back to the late 1960s, with very basic wire frame visualisations and ceiling-mounted arms to track head movements (Sutherland, 1968). Around the same time, studies using pseudo-participants (individuals pretending to be participants, who actually knew the goals and therefore acted in a specific way) were being explored (Asch, 1956; Latané & Rodin, 1969). The use of pseudo-participants does add a level of consistency to an ecologically valid experiment, compared to an observational study, however it does not match the level of uniformity and control that VR accounts for. An ideal experimental set-up would involve no extraneous variables, with only the variables that the researcher wishing to change actually doing so; in the real world this is near-impossible to create, even with the use of pseudo-participants. However, in virtual space this is much more attainable, and with features such as high-definition visual graphics and high fidelity sound, the realism in VR will be able to overcome the artificiality of simple display systems.

Experiments built in controlled environments with calibration measures in place are ideal for scientists to determine causation between variables and the impact specific factors have on one another. However, such laboratory experiments lack naturalistic viewing experiences. For example, we can speculate on the likelihood that findings of an attention task, when fully immersed and primed on the highly constrained task, will actually apply when walking down the street during a daily activity.

Conversely, in a free-flowing environment, having participants able to act naturally allows for the data collected to be more appropriately applied to everyday contexts. A study by Sanchez-Vives and Slater (2005) found that participants in VR behaved as if environment was real, which is why using VR provides data which has greater ecological validity than traditional computers. However, with the addition of naturalistic behaviour from participants comes the complication of having many additional metrics that may be recorded. For example, add-on devices can be paired with VR headsets for

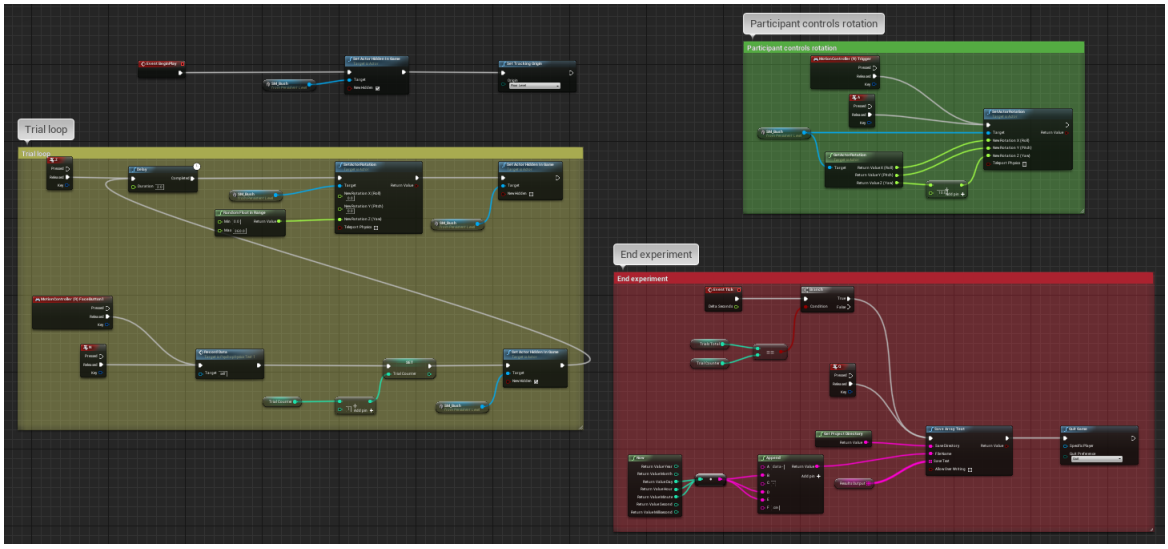


Figure 2.1: The complete level blueprint for the experiment to be created via this tutorial.

experiments which can record body movements and physiological responses to stimuli. Accessories such as the Leap Motion Controller allow for touchless interaction with virtual environments, and the STRATOS Inspire aid in touchless haptics (Ultraleap, 2021). Items such as these provide hand-tracking to enhance the subjective experience but also increase the scope for measurable metrics during experiments.

Having a free-flowing procedure in VR, while maintaining control over extraneous variables that may have an impact if tested in physical space, is a major benefit of VR over other experimental technologies and paradigms. Games engines allow for the creation of this type of world to interact with. This however, then brings the challenge of bringing sufficient structure and control to the environment, and the participant's interaction with it, to allow data collected in a way that allows standard psychophysical procedures of experiment design and data analysis to be applied. The rudimentary example developed in this chapter demonstrates how this can be achieved.

2.1.3 What Unreal Engine Is

Unreal allows you to access and control the power of VR for vision science. For those who are just starting to learn how to code, or do not have the resources to learn, or who

have a preference for visually building, blueprint scripting can be a useful alternative. Blueprint scripting is a visual approach to building an experiment or experience, where the execution flow consists of linearly connected nodes. Nodes are blueprint components such as events, function calls, or variables. UE uses a blueprint interface, rather than traditional scripts with lines of code, whereby reusable components can be inserted to the flowchart-like interface. Depicted in *Figure 2.1* is a screenshot from a basic `Level Blueprint`. Regardless of the application for the level, the building process of the blueprint is identical, whether it is intended for a mobile application, PC game, or research experiment. Building blueprints in a level-based format allows for precise organisation: levels can be conditions of an experiment, or actual levels in a game; they are simply separate categories for the end piece of software.

UE4 is undoubtedly the most used version of the engine because its release is recent but equally it has been out for enough time for its use to expand beyond previous versions; therefore, this version will be the focus of this article. Throughout the sections, relevant literature will be presented and a tutorial for creating an experiment with these specific parts will be explained. For any additional help for the programme, there is the the UE4 Documentation site with tips on how to navigate the programme, and within UE4 additional information can be provided on many of the functions by hovering over them with the mouse while pressing *Ctrl + Alt*.

UE also caters for programmers to incorporate C++ code into the editor, to accelerate development workflows. Changes to interface elements and editor functionality can be written into the blueprint system so that new and unique components can be created. One of the main aims of C++ is to support object-oriented programming (Stroustrup, 2007), which is why it fits perfectly with UE. It is possible to create new nodes, using C++ (see *Section 2.7.3*), and a using source code editor within a development environment such as Notepad++, Microsoft Visual Studio, or Apple's Xcode.

UE uses a left-handed Z-up co-ordinate system, in that the X axis is forward/back, Y is left/right, and Z axis is up/down. This is unique as most other games engines

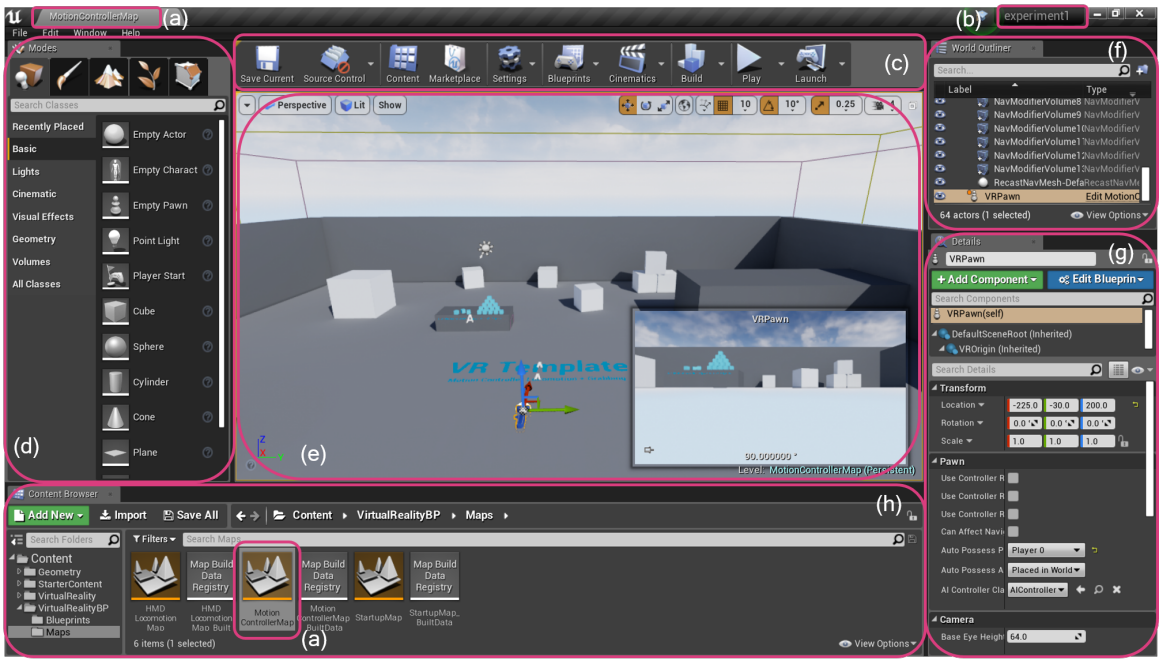


Figure 2.2: View of a starter level. (a) current level; (b) current project; (c) `Toolbar` window; (d) `Modes` window; (e) `Level Editor` Viewport; (f) `World Outliner` window; (g) `Details` window; (h) `Content Browser` window. Axes shown in corner of Level Editor, with X corresponding to location (distance, forwards/backwards); Y also to location (left/right); Z to height (up and down).

and software packages use either the right-handed Y-up (including Maya, Godot), right-handed Z-up (Blender, Cryengine), or left-handed Y-up (Unity, ZBrush).

2.1.4 Applications of Unreal Engine

UE was first developed to build the first-person shooter *Unreal* (Thomsen, 2012). Since then it has grown in power and popularity and produced large amounts of software applications. The focus on providing a streamlined workflow for artists and designers to work with developing is evident in the blueprint style format of the engine. A number of the types of programs developers have used it for are discussed here.

The rooting of UE is presentation on 2D screens: environments with limited degrees of freedom such as *Tetris Effect* (Monstars Inc. and Resonair, 2019). In this revamp of the 1980's classic arcade game (Hoad, 2014) game players adjust the roll axis of falling tetrominoes aiming for them to be able to fall into vacant spaces at the bottom of the screen. In contrast, UE has the capabilities to create much more complex games, such

as *Borderlands 3* (Gearbox Software, 2019): atmospheric lighting effects, sub-surface scattering, and the unique outline methods of presenting objects in the game franchise, were programmed in UE4 (Strickland, 2017). Additionally, UE4 was chosen for the creation of this game as it allows the end product to be ported onto an array of different platforms (McKeand, 2019).

Other platforms that UE has the potential to port onto are mobile phones, including both Apple iOS and Android operating systems: UE3 was used to create *Mortal Kombat X* (NetherRealm Studios, 2015). Even in PlayStation VR the standalone game *Batman: Arkham VR* utilises the first-person perspective to play through the eyes of the playable character, Batman (Wawro, 2016).

Unreal Engine also allows for the development of VR environments. An immersive VR experience created by Unreal Showdown Cinematic VR Experience (Ivey, 2015) showcased many of the capabilities and was the the start of many more vicarious cinematic experiences. A collection of immersive environments created by the Savannah College of Art and Design, named Destination Imagination has used UE4 to take users on a 'thrilling' journey into outer space (Robinson, 2018). The premise of the experience is that the user has escaped the massive destruction of planet Earth and the observable scenes are the processes of becoming an astronaut while leaving the atmosphere Defreitas (2018). Robinson reported that the creators chose UE4 to create this project because of the approachability and power of the engine, where it is possible to quickly test and adjust ideas and having a high fidelity and lighting structures ready to manipulate.

While all of these games and experiences are commercially available either for free download or purchase, the techniques that developers used to create them can be implemented into the field of research. Using UE to create environments on 3D computer screens, perhaps for psychophysical experiments, is achievable; it is just much less developed, and shared, with the vision research community at this point in time than the established platforms of MATLAB and Python.

A brief search through recently published papers uncovered very few which use UE.

Previous research conducted into depth perception by the author has shown its usefulness in collecting the desired metrics for quantitative experimentation (*Chapters 3, 4, 5*), hence this paper describing the uses of UE and why it is an appropriate program for such research.

2.1.5 Key for This Guide

To access UE, the Epic Games Launcher needs to first be downloaded onto the computer being used. In the current tutorial there are steps to follow which are in typewriter font and often centred. Each step will be separated by a directional greater than symbol. Adjustments to be made by the reader are described within asterisks, often indicating a visible box should be ticked, a button to be selected, or asset to be renamed. An example is as follows:

`step 1 > step 2 > *information for reader* > step 3`

Instructions for the tutorial side of this article will be labelled as such, and encased in a pink box. By following the instructions which are presented in this way, a fully functioning VR experiment will be created. The tutorial was built using version 4.20.3 of UE, if following this using a different version there may be slight differences.

There are some instances where a program such as Microsoft Visual Studio will be necessary to add specific coding, it is suggested that this is installed before starting the tutorial. Errors, especially when it comes to that presented in *Section 2.7.3* can be fixed or even bypassed by ensuring all updates for this are installed. UE has been designed to be integrated with Visual Studio, when installing the software ensure that `Game development with C++` is selected in the initial pop-up, and that the individual components of `C++ profiling tools, Windows 10 SDK, Unreal Engine installer` are selected when installing. For more tips when installing this, the Unreal Documentation provides an extensive explanation.

SteamVR should also be installed, and signed in, for the VR headset to work with UE4. This can be downloaded from the Steam website where tips are also provided.

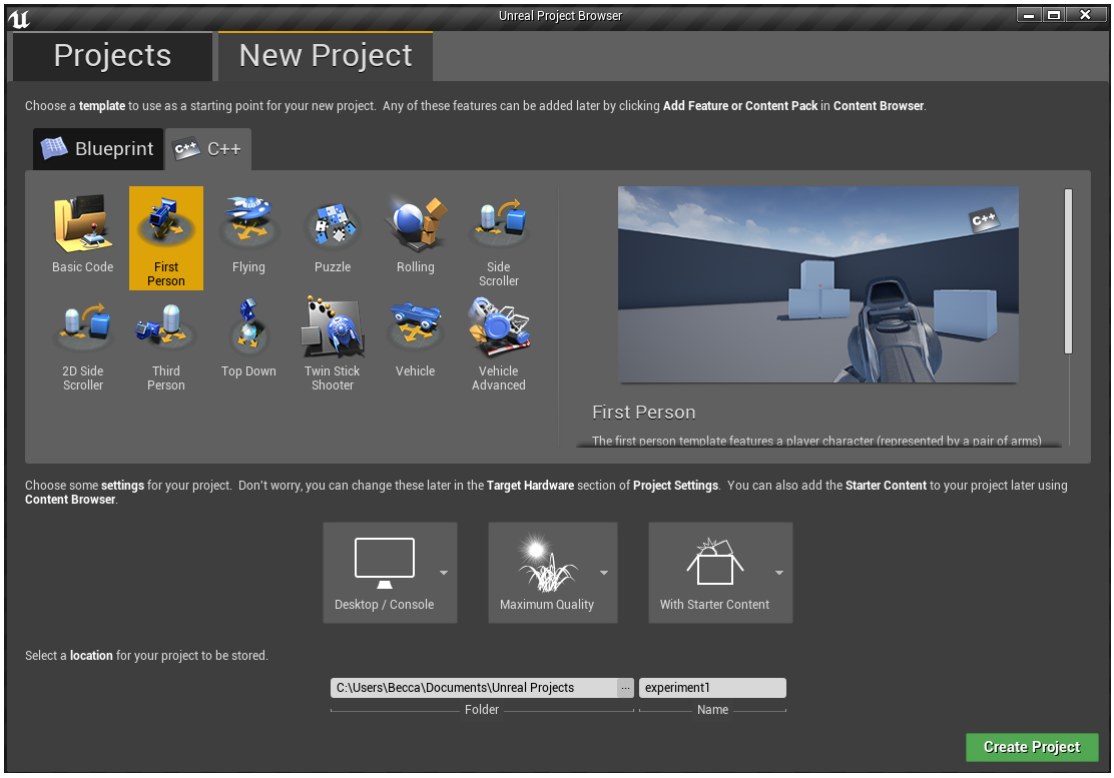


Figure 2.3: View of the Unreal Project Browser, where a new C++, first person project is to be created. The project will be for a Desktop at maximum quality, contain starter content, and be named 'experiment1'.

This is needed to link the headset with the UE software. Following this, right-click on the SteamVR window and select the 'Run Room Setup' and follow the on-screen instructions to set up the SteamVR interaction area.

Before starting this tutorial, it is also important that the required VR headset is set-up and running on the PC being used. If the tutorial is followed, a blueprint such as that shown in *Figure 2.1* will be created.

2.2 Creating an Environment

2.2.1 Starting a Project

An experiment can contain many different conditions; when using UE different experiments can be created in the form of projects, and conditions may be organised as levels within these projects.

In this guide, the project is named "experiment1" which is to be on a Desktop and loaded with Starter Content. Once created, the organisation of the windows shown in *Figure 2.2* can be rearranged and resized, for personal preference. The `World Outliner` is the window that displays all assets inserted into a level, which can be organised in folders by clicking the folder icon in the right window. There are a number of example levels pre-loaded into this blueprint which can be deleted to reduce the size of the saved game, however they are not visible to the end user if they remain.

Instruction: create project

To create a new project, launch the Epic Games application and follow these steps to create a new Virtual Reality Blueprint, as displayed in *Figure 2.3*:

```
Unreal Engine > Launch Unreal Engine > New Project > C++ > First
      Person > *select 'Starter Content'* > *name project
               ("experiment1")* > Create Project
```

The instructions in this guide are organised in such a way that C++ coding can be added (*Section 2.7.3*). However, if there is no desired output of the project, for instance is a virtual experience is to be created, then the following can be followed as an alternative:

```
Unreal Engine > Launch Unreal Engine > New Project > Blueprint > Virtual
Reality > *select 'Starter Content'* > *name project ("experiment1")* >
                          Create Project
```

From here, three maps are included as standard, delete the `HDLocomotionMap` and `StartupMap`); only the `MotionControllerMap` should be retained. In the `World Outliner` it is possible to delete the contents of the environment, by selecting and pressing 'Delete' on the keyboard: do this for everything except the `MotionControllerMap` and `VRPawn`; if the latter is accidentally deleted, this can be retrieved from:

```
Modes > All Classes > *search "MotionControllerPawn"* > *drag and drop
                          into the level*
```

The rest of the instructions can be followed from here, however as this project has

the VR capabilities already set-up there will be no need to set up the controllers.

New levels within a project can be created via `*right-click in Content Browser* > Create Basic Asset > Level`. Alternatively, one of the ready-made levels can be cleared of assets (from the `World Outliner` window, select all then press 'Delete' on the keyboard). Creating different levels to be different conditions is one way to organise a project.

Instruction: format level

Delete the three loaded levels in `Content Browser > Content > StarterContent > Maps`, navigate to `Content Browser > Content > FirstPersonCPP > Maps`. Delete the remaining map. Create new folder for everything relating to this experiment, and create a blank level.

```
Content Browser > *right-click on empty space* > New Folder >
*rename folder ("ForExperiment")* > *right-click on empty space* >
Create Basic Asset > Level > *rename level ("PsychophysicsTest1")*
```

Make sure to drag and drop `PsychophysicsTest1` into this folder if it is created in a different folder. Set this level as the default when this project is opened.

```
Edit > Project Settings > Project > Maps & Modes > Default Modes >
  Default GameMode > *select 'GameMode'* > Default Maps > Editor
  Startup Map > *select 'PsychophysicsTest1' > Game Default Map >
                  *select 'PsychophysicsTest1'*
```

In many experiments, participants are required to use controllers, or physically move. In UE4 it is necessary to program the exact button presses into the blueprint, and create the output for what the button press means. A button press can be used as a response to a stimulus, so the outcome of this press would not cause a visible effect; or a participant may be required to press buttons to adjust aspects of the environment. Setting up controllers is described in *Section 2.6.1*, however sometimes only the virtual representation of the participant (also referred to as an *avatar*) needs to be formatted into the experiment. The motion controller pawn is set up for the use of VR controllers: in *Figure 2.2* the `Motion Controller Map` is presented, with the VR Pawn selected.

This pawn represents the VR user: wherever the participant should be positioned in the environment can be manipulated via this pawn. On the other hand, if a C++ Controllers project was created (as per the instructions) then continuing to follow the tutorial will allow for this to be set up.

The `Player Start` class should always be added to the environment and possessed as 'Player 0', as this is where the participant will appear when they enter the experiment. Depending on the perspective required, the location across the X and Y axes can be adjusted. A few additional teaks should be made to ensure this syncs with the VR headset:

Instruction: format participant position

Pick a location for the participant to appear at on startup.

```
Modes > Player Start > *drag and drop into the level* > Details >
Transform > *adjust axis positions to X = -70, Y = 0, Z = 100* >
Modes > *search and drag 'Camera' into the level* > *adjust axis
  positions to = -70, Y = 0, Z = 100* > World Outliner > *drag
      'Camera Actor' and drop on top of 'Player Start'*
```

Sync the player start with the VR headset.

```
*ensure 'PlayerStart' is selected > Details > Input > Auto Receive
                  Input > *select 'Player 0'*
```

Allow the VR headset to be used during preview and gameplay.

```
   Edit > Project Settings > Project > Description > Settings >
                    *select 'start in VR'*
```

All levels, materials, assets, etc can be renamed in both the `Content Browser`, and also the `World Outliner` for the specific versions of the assets placed into the level. Folders can even be created in the `World Outliner` by right-clicking here: dragging and dropping similar groupings of items into one folder helps with tidiness and makes it easier to quickly find specific assets.

It is good practise to frequently save the project (`Toolbar > Save Current`) as well as the whole project (`File > Save All Levels`), as well as using `Build` in the Toolbar by pressing it before previewing to pre-compute the lighting and visibility data.

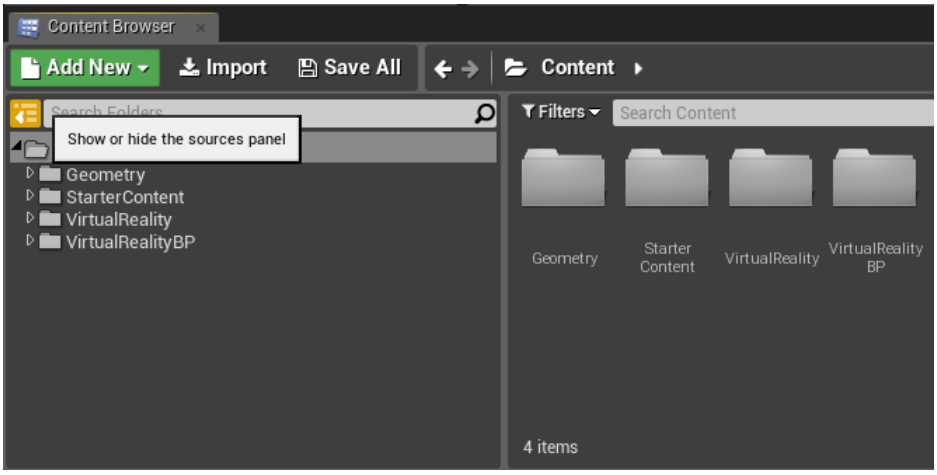


Figure 2.4: Screenshot of the `Content Browser`, where the highlighted button allows for an alternative view of the content folders.

When navigating around UE4, mouse movements are key: while holding either the left mouse button (rotate and move in any lateral direction) or right button (rotate across all angles) in the `Level Editor` you can adjust your viewing of the environment you have created. Holding the right button in a blueprint will allow you to adjust the area of the blueprint that is viewed. The scroller can also be used to zoom in and out.

Accessing the blueprints for different levels can be found via `Toolbar > Blueprints > Open Level Blueprint`. When making adjustments to the settings in pop-up windows, there is no button to save the adjustments; when the close button is pressed, the settings are automatically saved.

Navigating through the different folders for the maps or assets can be made quicker by opening the sources panel, as depicted in *Figure 2.4*. This causes all folders to be visible in their appropriate structure.

The top left corner of the `Content Editor` (*Figure 2.2 (e)*) has four buttons: the first drop-down menu (downward arrow) consists of viewport options, where different layouts and setting of the programme itself can be altered. The second drop-down menu (where 'Perspective' is currently selected) allows for different viewing options of the virtual environment itself, changing the viewpoint for the creator to observe the

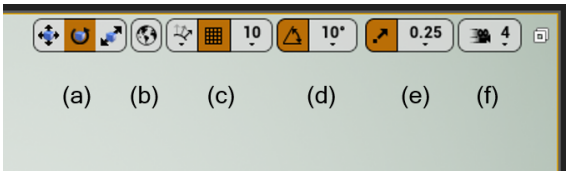


Figure 2.5: Screenshot of the top right corner of the `Content Editor`, where grid options are available. Group (a) is the current transformation settings, the three options are moving, rotating, sizing. (b) is the Cycle Transform Gizmo, where the axis of an object in the scene can be altered, depending on the direction it is facing in the environment. Group (c) are the snapping options, when moving an object in the scene it is possible to get it to snap to specified increments, or not. Group (d) is the same for snapping the rotation movements. Group (e) is the same for snapping the size scaling increments. Group (f) is the camera movement speed.

scenes from different perspectives. The third (where 'Lit' is currently selected) allows for different observations of the lighting conditions. The final drop-down menu ('Show') is where the creator can choose what specific aspects of the scene are visible. None of these settings need to be altered for the tutorial side of this chapter.

The top right corner of the `Content Editor` (*Figure 2.2 (e)*) has six groups of buttons, a close-up is shown in *Figure 2.5*. These groups of buttons give greater control over the way objects can be positioned in the virtual environment: snapping options ensure that an object can be aligned accurately (a floor and a wall, for example), and the increments of centimetre (translation), degrees (rotation), or percentage (scaling) that objects can be adjusted by.

2.2.2 Obtaining Assets

When starting a new project, one of the easiest ways to create an environment is to use the Starter Content within the `Content Browser`. Along with 3D objects (window, chair, tree, to name a few) it also imports materials and textures to enhance the realism of these structures.

In the content builder it is possible to create new structures from the pre-loaded content: a 'quick and dirty' way to create the large table shown in *Figure 2.6* is to rotate and move a wall as the surface, and two door frames as the legs. This table was created for use in *Chapter 5* because the method required a long table with specific



Figure 2.6: Screenshot of a table made from separate pieces of Starter Content, where the selected front legs are made from a door frame.

measurements in order to conduct the equi-distant task. Often it is quicker to make a simple object rather than spend time and money searching and purchasing a pre-made asset which may not meet the requirements of the experiment.

Unreal Marketplace is the primary source recommended for developers to obtain new content to add into projects. Here it is possible to get blueprints for objects which have been created by artists or scanned and uploaded by developers. There are many search options to narrow down the search of any asset desired. When one has been added to your basket and checked out, it is then easy to select it in the Epic Games Launcher marketplace where you can add it to any previously created project. TurboSquid is another alternative online store, which has a large range of texture maps in addition to 3D models.

It is important to bear in mind that UE uses .fbx files for its assets, so when obtain-

ing or creating assets to import into the program it will need to be in this format. UE works well to reformat other file types, such as .obj, however this can often cause it to crash when the file type is too large. In these cases, simplification and reformatting can be performed using an external program, such as Meshlab.

To create a virtual representation of the VR controller being used, it is likely that one has been installed on your device when SteamVR was installed at the start of this tutorial. In the file explorer, a range of controller meshes can be found at:

`C: \Program Files (x86) \Steam \steamapps \common \SteamVR \resources \rendermodels`

From here, there is likely to be a folder associated with the VR headset that is being used. If the HTC Vive is being used, for example, by navigating through the above navigation, then to the `vr_controller_vive_1_5` folder, and selecting the following: `onepointfive_occ_bake.tga` (for occlusion bake), `onepointfive_spec.png` (for specular information), `onepointfive_texture.png` (texture), `vr_controller_vive_1_5.mtl` (material), `vr_controller_vive_1_5.obj` (object mesh), these will create a virtual representation for the Vive controller. Depending on the version of UE, the material file may not upload, in which case this material can be created based on the other files uploaded.

A pop-up box will appear, names `FBX Import Options`: the `Import Uniform Scale` should be adjusted to 100, before importing all. In the `Content Browser` these files should all now appear, double click The .obj (which has now been converted to a static mesh) file so that the material can be adjusted:

`Details > LOD0 > Element 0 > *double-click the white circle which is labelled 'whole_model_group1Mat' to open a new blueprint* > *delete 'None Param(0.8,0.8,0.8)' > *on keyboard hold 'T' and left click* > Details > Material Expression Texture Base > Texture > *select from drop-down menu 'onepointfive_texture'* > *connect white output pin 'Texture Sample' to 'whole_model_group1Mat Base Color'* > *on keyboard hold 'T' and left click* > Details > Material Expression Texture Base > Texture > *select from drop-down menu 'onepointfive_spec'* > *connect white output pin`

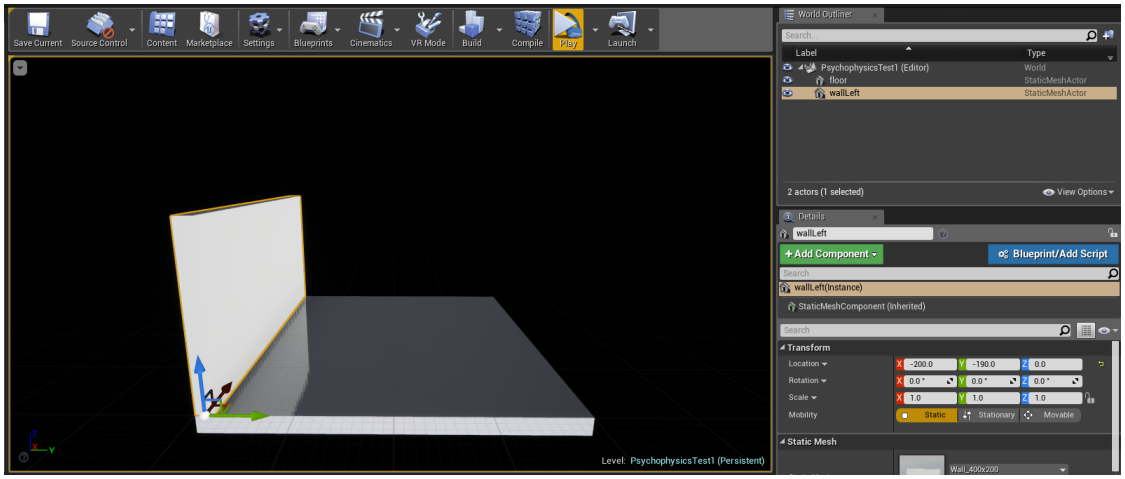


Figure 2.7: Creating an environment, the 'wallLeft' is selected in `World Outliner`, with location information sown in the `Details` window.

`'Texture Sample' to 'whole_model_group1Mat Specular'* > *on keyboard hold 'T' and left click* > Details > Material Expression Texture Base > Texture > *select from drop-down menu 'onepointfive_occ_bake'* > *connect white output pin 'Texture Sample' to 'whole_model_group1Mat Ambient Occlusion'*`

This can then be added as an avatar for the observer to see when in the virtual environment, and when holding the controller (see *Section 2.6.1*) if this is desired over some other representation of the hands.

2.2.3 Positioning Assets

The measurements used in UE vary depending on the version being used: in UE3 the units are Unreal Units (uu), where one uu is equal to 0.75 inches but in UE4 however, 1 uu is equal to 1 cm by default (Emperore & Sherry, 2015). It is important to know this when importing older projects into the new version as this is likely to affect the findings of any experiments being re-run on a newer version.

Inserting an object into the environment before the experiment begins, during the creation phase, is simple. Walls, ceiling, and a floor can all be inserted into the `Level Outliner`, and positioned in such as way as to create a room for the experiment to be conducted in. Once inserted they can then either be dragged to the desired location

in the `Level Viewer`, or a specific location for the X, Y, and Z axes can be specified. These assets can be duplicated for ease, which streamlines the process if certain features are manipulated (such as size or material) and are needed to be repeated.

Instruction: create room

Create a floor, as depicted in *Figure 2.7*.

```
Content Browser > Starter Content > Architecture > *drag
Floor_400x400 into Level Editor > World Outliner > *rename
("floor")* > Details > Transform > *set the Z axis location to 0,
set X and Y location to -200*
```

Create four walls, repeat this for a left, right, front, back wall.

```
Content Browser > Starter Content > *drag Wall_400x200 into Level
Editor > *rename ("wallLeft")* > *drag to the edge of the floor* >
*press 'E' on the keyboard to allow rotating if necessary; press W
to switch back to moving*
```

Create a ceiling.

```
*select the floor* > *on keyboard press 'Ctrl W'* > *rename
("ceiling")* > Details > Transform > *adjust to same XY axes as the
floor, Z axis to the height of the walls = 200*
```

It is good practice to centre the ground level surface in an experiment around the Z axis location of 0. In *Figure 2.7*, the Ceiling is a duplicate of Floor, in order to move it to the appropriate position the Z axis should be adjusted to the height of the walls (200). The same process can be used to rotate the Z axis of any walls and other aspects of the scene which intend on being positioned in set areas. A shortcut to switch between rotating and translating is by pressing 'E' on the keyboard to use mouse movements to rotate in the `World Outliner`, and 'W' to switch back to moving the object. The same shortcut applies to scaling the size by pressing 'R' on the keyboard.

At the top right corner of the World View there are three drop-down options which are useful when positioning objects in the environment: the first is Surface Snapping. When this is toggled on, any assets that are being positioned in the environment will

automatically sit on any floor plane which is present. This is very useful when creating a naturalistic scene as it negates the need to consistently run the application to check if objects are positioned on the surface. In some experiments, however, it is necessary that the stimuli are not positioned on a surface, or that they are in a very specific location.

Ensuring a stimulus is presented at the exact location is vital in many experiments. As a typical example a study by Sprague et al. (2016) required a visual stimulus to be placed 115 cm in front of the observer. The blur on the stimuli was manipulated between trials to test whether perception of depth within the image arose from vertical or horizontal blur gradients. The findings suggested participants tended to perceived sharper areas of an image as being closer, and areas with more blur were perceived as being further away. In this example, it was necessary to maintain a constant distance to the actual stimuli, so that any perceived change in distance or depth could be concluded as being a result of the blur being varied.

Instruction: insert stimulus

Add an object from the Starter Content to act as the stimulus in the experiment

```
Content Browser > Starter Content > Props > *drag 'Stimulus' into
the Level Editor* > *rename ("Stimulus") > Details > Transform >
            *set location:  X = 120, Y = 0, Z = 0.75*
```

2.2.4 Asset Materials

Changing the appearance of any asset placed in the project is a simple case of drag and drop for any materials or textures pre-loaded in the `Content Browser`, onto the asset itself. UE4 will automatically convert a texture into a material if done this way. Materials can be found in all of the locations described in the *Section 2.2.2*, but creating unique materials is also simple. By dragging and dropping an image file into the `Content Browser` (or loading it in manually) UE4 automatically converts the texture into a usable material to be placed within the environment. By double clicking on the material, edits to specific components of the material can be made, such as the specularity and emissive

effects.

If any of the pre-loaded content from the Starter Content is inserted into the environment, the associated material and texture will automatically be set.

An example of when the appearance of stimuli has needed to be adjusted during a psychophysical experiment can be found in the discrimination experiments by Sharman et al. (2013). The method required the photographs of stimuli to be well lit and in focus; he stimuli were then manipulated to vary in their blur/sharpness and chromaticity/luminance so that the interactions between sharp luminance information and blurred chromatic information could be investigated in natural scenes. In UE, there are similar options to adjust many other visual settings so that the materials can appear to be made from different substances such as liquid, metal, or fluorescent lighting.

Instruction: adjust room appearance

Drag and drop materials onto the surfaces of the room, as shown in *Figure 2.8*.

`Content Browser > Starter Content > Materials > *drag M_Wood_Floor_Walnut_Worn onto the floor in Level Editor* > Content Browser > Starter Content > Materials > *drag M_Concrete_Poured onto the ceiling in Level Editor*`

Materials of objects can be set via two ways, which is depicted in *Figure 2.8*. In *(a)*, the material `M_Wood_Floor_Walnut_Worn` has been selected is being dragged to the floor in the `Level Editor`. An alternative way is to select the wall from the `World Outliner`, scroll to `Materials` in `Details` and select the desired material from the drop-down menu, also shown in *(b)*.

2.2.5 Lighting

Four types of lighting can be used, from the `Modes` panel: `Directional Light, Point Light, Spot Light, Sky Light`. If an ambient light source which will not be later manipulated during the experiment is desired, after dragging one into the scene the



Figure 2.8: (a) Floor material being dragged and dropped to surface. (b) Ceiling material being adjusted via Details window.

`Mobility` of the light in `Details` should be set to `Static`, so that it renders quickly. Many aspects of the lighting can be manipulated through UE4, and there are even more options being released in UE5. Two features to be covered here will be the luminance and colour.

Lumens are a measure of the total quantity of visible light emitted by a source, per unit of time. It may be that Lux is a more familiar unit, in which case the conversion is one lux being equal to one lumen per square meter. The default unit of UE4 is Candelas, which is equal to the luminous flux (in lumens) divided by the solid angle (in steradians). The unit used can be adjusted between unitless, lumens, or candelas via:

`Edit > Project Settings > Engine > Rendering > Default Settings > Light Units > *select appropriate option*`

The light colour can also be an important feature. In a study by Iriguchi et al. (2018), the colours of three cartoon faces were used as stimuli. The colour options were red, green, and blue and the aim was to test whether going through menopause contributed to colour perception. The lit stimulus colour was randomised between trials; reaction time and response were recorded. It was found that pre-menopausal women had faster reaction times than post-menopausal women, and that overall blue was the slowest colour to be identified. Similarly, Hibbard et al. (2017b) tested the effect of differing luminance on depth judgements in stereoscopic displays and found luminance did have an impact on binocular depth perception.

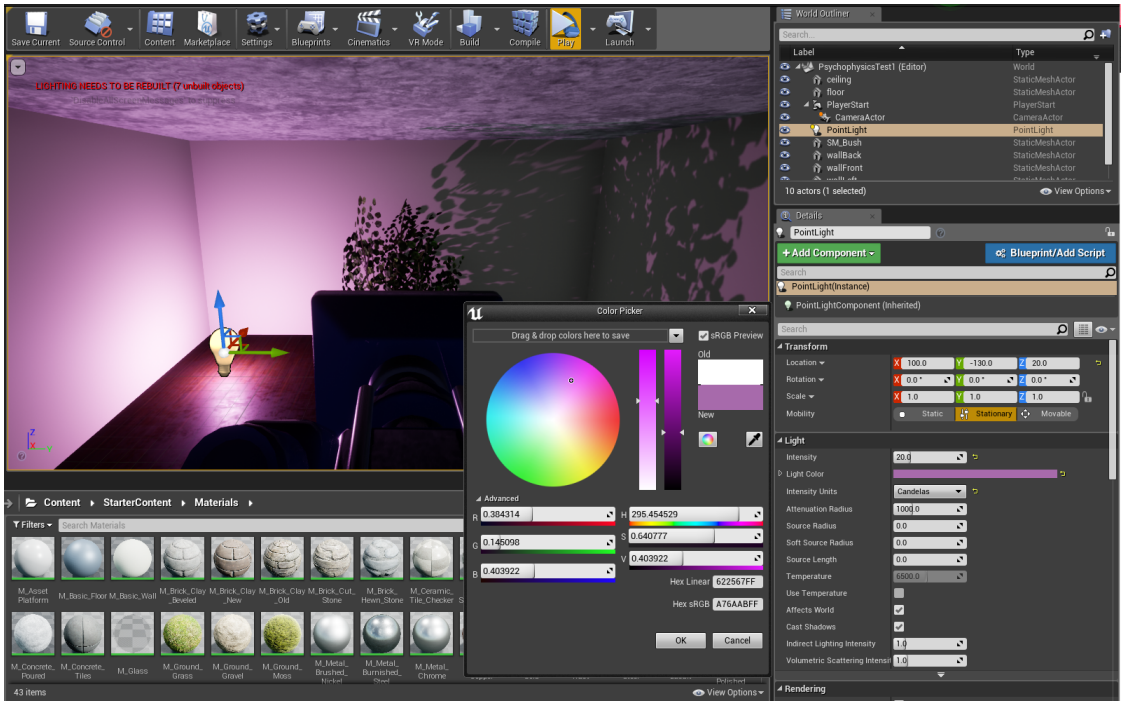


Figure 2.9: Colour picker in Details window can be used to adjust the colour of lights.

In the `Details` panel, the colour properties such as saturation and hue of any light can be adjusted. The colour picker in *Figure 2.9* depicts the colour wheel which appears in the `Details` window once a light is inserted into the level; this presents possible colour options to choose from, but the RGB and Hex codes can also be used by way of adjustments.

When making adjustments to the lighting, it is important to build the lighting into the project (`Toolbar > Build`), as well as saving and compiling.

Instruction: add lighting

Add a light source so that participants can see when in the environment.

```
Modes > Lights > *drag Point Light into the Level Editor* > Details
  > Transform > *set location axes to X = 100, Y = -130, Z = 20 >
                Light > *set Intensity to 20*
```

2.3 Stimuli

2.3.1 Sound Effects

Sound is frequently used to initiating the start of a trial or even representing the the stimulus itself (Hanley & Bourgaize, 2018; Jonas et al., 2017). Any pre-recorded sound effect can be imported into UE4 via selecting `Import` in the `Content Browser` and selecting the saved file. As with importing most things into UE4, it is also possible to drag and drop the file from any file explorer into the `Content Browser`.

The same steps can be taken to import a computer-generated noise, for example MATLAB has a useful function for creating a single-tone chime with any set duration time (Star Strider, 2015). UE4 can even be used to create a custom sound cue, where factors such as the distance range and occlusion can be manipulated from the `Content Browser`:

```
Add New > Sounds > Sound Cue
```

It may be necessary to sync up the initiation of a sound effect to a timeline, in which case *Section 2.4.3* should be used in conjunction with this information.

2.3.2 Moving Objects

Paradigms can involve automatic movement of the stimuli. Dyde et al. (2011) conducted a study whereby recognition of image silhouettes was tested: in each trial, the stimulus was presented at differing rotations between 90 to 270°. The experiment was set up to automatically, and randomly present one of the image variants in each trial. Reaction times between presentation and a button press from the participant were recorded; it was found that reaction time was influenced by image orientation, and the orientation also biased the perceived direction of hand (left or right).

In contrast to this, Kline and Witmer (1996) conducted an experiment during which the stimulus (a wall) was automatically moved between trials. Factors affecting distance estimation were being examined, the task required participants to give estimates on the

amount of space separating them from the wall. It was found that field of view within a VR headset affected distance perception more than textures on the stimulus wall did.

In the `Level Blueprint`, the `Get Actor Location` node can be used to retrieve the current location of an object, and `Set Actor Location` node can be used to transform the location of an object. In *Figure 2.10* (taken from *Chapter 5*) the locations of the `Reference` and `Target` are randomised to be a random position between 100 and 300 cm along the X axis when the *Q* key is pressed. Much like the above pieces of research, it is possible to get an object to move at a specific time, as detailed in *Section 2.4.2*.

Here, we can format the project so that when the experiment begins, the stimulus will randomly rotate an amount after a delay. At the start of each trial this will also occur.

Instruction: create start of trial

Create experimenter input to start experiment, followed by 2 second delay, stimulus appears and rotates an amount. This will create the first line of nodes in *Figure 2.11*.

```
World Outliner > *ensure 'Stimulus' is selected* > Toolbar >
Blueprints > Open Level Blueprint > *right-click empty space on
blueprint* > *search and select 'Z' from 'Keyboard Events'* > *drag
from 'Z Released Exec'* > *search and select 'Delay'* > *adjust
'Delay Duration' to "2"* > *right-click empty space on blueprint* >
Create a Reference to Stimulus > *drag output of 'Stimulus
Reference'* > *search and select 'SetActorRotation'* > *connect
'Delay Completed Exec' to 'SetActorRotation Exec'* > *right-click
'New Rotation' on 'SetActorRotation'* > Split Struct Pin > *drag
from 'New Rotation Z (Yaw)* > *search and select 'Random Float in
Range'* > *set 'Random Float in Range Max' to "360"* > *drag from
'SetActorRotation Exec', release mouse click'* > *search 'visible',
select 'Set Actor Hidden In Game'* > *right-click empty space on
blueprint* > Create a Reference to Stimulus > *link output of
'Stimulus Reference' to 'Set Actor Hidden In Game Target'*
```

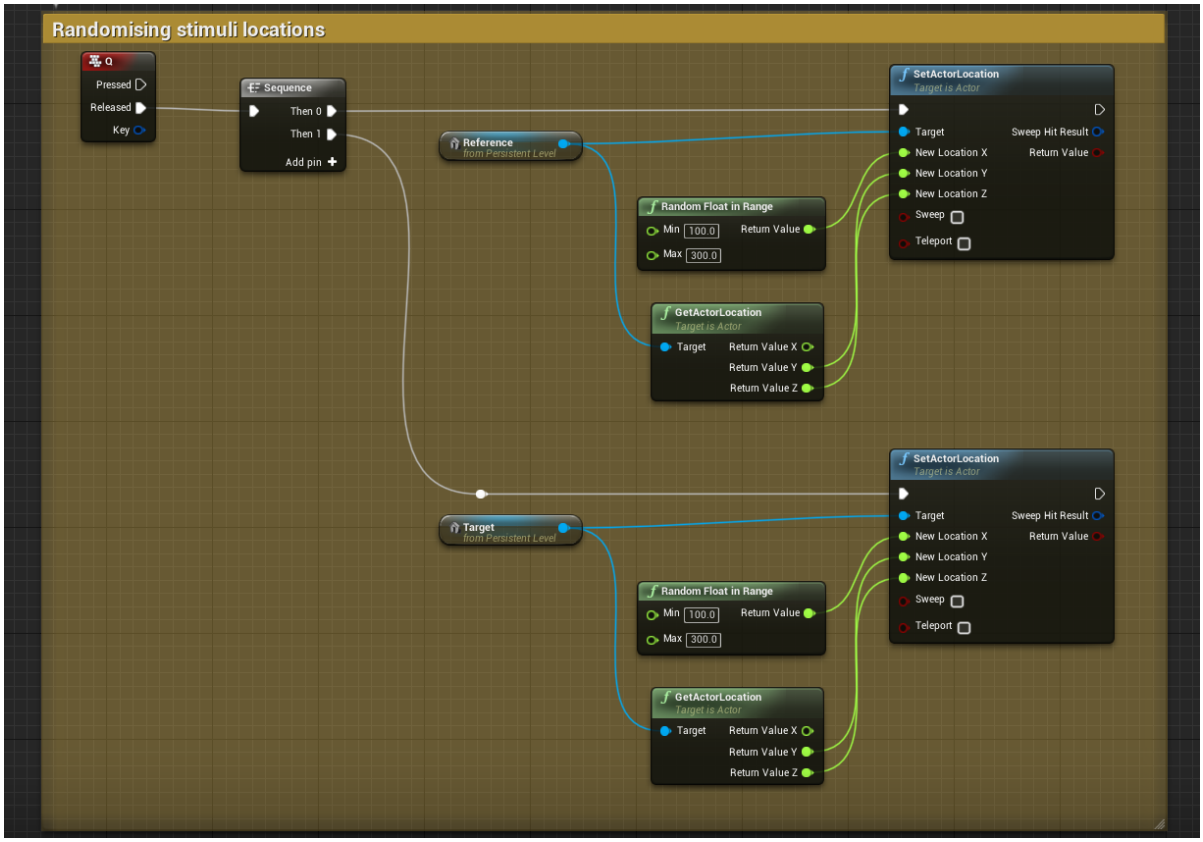


Figure 2.10: When 'Q' on the keyboard is pressed, two actors labelled Reference and Target are randomly positioned at locations between 100 to 300 along the X axis. These nodes have been selected together and commented as "Randomising stimuli locations".

2.3.3 Spawning New Objects

When an object is referred to as being *spawned*, it means it appears in the environment, when it was otherwise not there. Some paradigms require the location of stimuli to alter between trials (*Chapter 3*) while others require slightly different stimuli to be presented between trials (Sharman et al., 2015). In UE, it is possible to spawn completely new assets within the experiment, as long as the asset being spawned is saved as a blueprint.

To set an object as a blueprint, let's take the `SM_Bush` static mesh actor from the `Starter Content` for example. After positioning it in the environment and making a note of where the desired appearance of it will be, in the `Details` panel there is a blue button named `Blueprint/Add Script`. When this is pressed, the name of the object which will now have a blueprint associated with it can be named and put into a folder for ease of access. The instance of this in the `World Outliner` can be deleted, as the

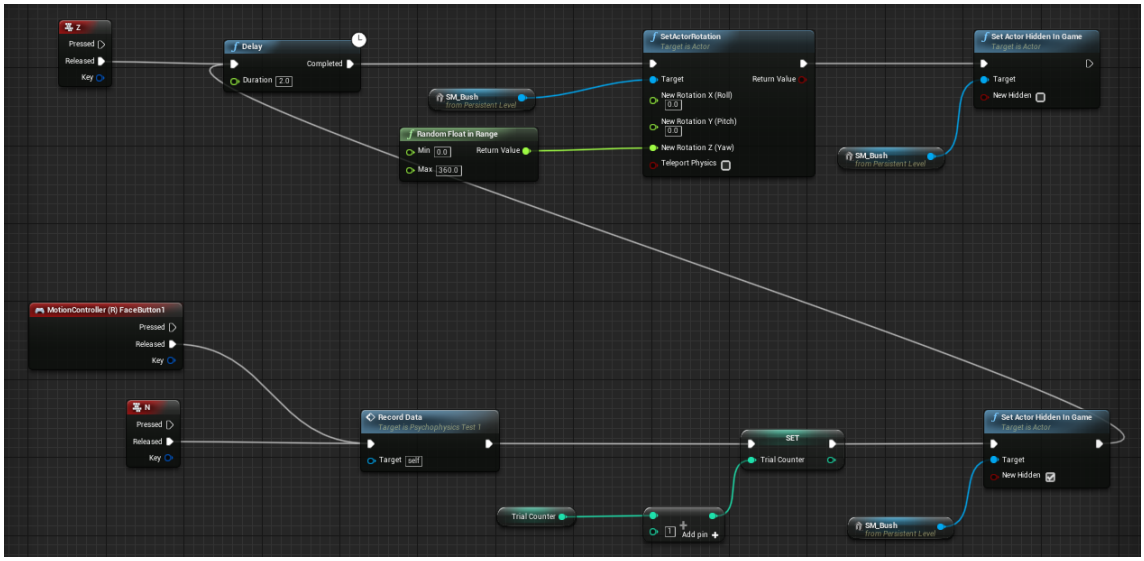


Figure 2.11: Blueprint depicting the loop that each trial goes through in the experiment.

blueprint version is now saved in the `Content Browser`.

Once this object has a blueprint we can turn to the `Level Blueprint` which can be accessed through `Blueprints > Open Level Blueprint`. By right-clicking and searching for the node `Spawn Actor from Class`, the blueprint can then be selected in the drop-down `Class` menu, and the transformation settings for which the object will appear at can be adjusted by right-clicking on `Spawn Transform > Split Struct Pin` and inserting the desired information. An input should be attached to this, perhaps, by pressing a key on a keyboard this object should appear: details of this can be found in *Section 2.4.2*.

On the other hand, it might be that an object is required to disappear mid-way through the experiment instead of being spawned. Either a static mesh or blueprint class version of the asset can be made to be invisible during parts of the experiment, once the object is positioned in the environment, switch to the `Level Blueprint` and insert the `Set Actor Hidden In Game` node. From this, a reference to your object can be attached to the 'Target' pin, and the Boolean tick-box of 'New Hidden' can either be selected or deselected. An input to this can be set-up (*Section 2.4.2*) and the output can either be linked to another event, or left blank. By doing this with any chosen

static mesh (or blueprint, let's take the original `SM_Bush` static mesh actor), it can be positioned and renamed to "Stimulus" in order to quickly understand that this is the stimulus which will be (dis)appearing throughout the experiment. If this is the case, the mobility of the asset needs to be set to `Movable`, so that aspects of it can actually be manipulated during gameplay.

Instruction: make stimulus not visible before start

When the participant enters the experiment, the stimulus will not be visible.

```
*ensure 'Stimulus' is selected* > Toolbar > Blueprints > Open Level
   Blueprint > *right-click on empty space* > *search and select
    'Event BeginPlay'* > *right-click on empty space* > *search
    'visible', select Set Actor Hidden In Game* > *drag 'Event
BeginPlay Exec' to 'Set Actor Hidden In Game Exec'* > *select 'Set
Actor Hidden In Game New Hidden'* > *drag from 'Set Actor Hidden In
   Game', release mouse click* > *search and select 'Set Tracking
    Origin'* > *select 'Set Tracking Origin Origin Floor Level'*
```

Ensure the stimulus is capable of being adjusted in this way during the experiment.

```
Level Blueprint > World Outliner > SM_Bush > Details > Transform >
                  Mobility > *select 'Movable'*
```

2.3.4 Changing Appearances

A stimulus which is required to change appearance within the experiment or game play is referred to as having a *dynamic material*. If it is only one object which is changing then it is one *instance* of this material. The ability to manipulate this is necessary if testing a series of variations of a texture; in a study by Sinai et al. (1999) for example, the floor material was altered between trials so that high (grass), medium (brick), and low (carpet) density textures could be examined for its effect on distance perception.

In UE4, an asset can be coded in such a way so as to have a changing material at a specific time, or after an input. In *Figure 2.12* the material of the `Wall_Back` actor is altered to be `M_CobbleStone_Smooth` when 'P' on the keyboard is pressed. Additionally, this can be set up in such a way as to alter the material of all actors in a specific class,

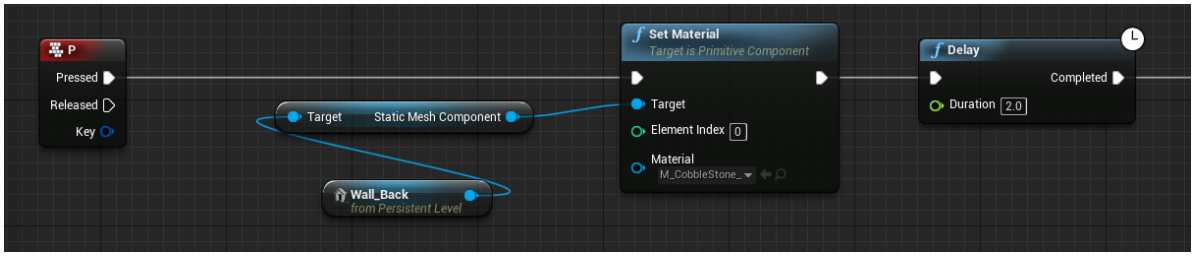


Figure 2.12: Section of a blueprint which changes the material of one object (named: Wall_Back), followed by a delay of two seconds before moving on.

if for example there are multiple stimuli presented to a participant where they all need to have a change in opaqueness instigated. Any material parameter can be selected to be adjusted; the way this would be coded for all objects within the same class would be to connect the `Get All Actors of Class` node to the Target of `Set Material` and selecting the appropriate actor class. Following on from the `Set Material` node, any other executables can be added, for example a delay (as pictured, *Section 2.4.5*) or any of the other aforementioned happenings such as the stimulus being set to disappear (*Section 2.3.3*).

Another particularly useful way to change the appearance of a stimulus is to create a dynamic instance and have specific parameters change on a timeline, as seen in *Figure 2.13*. By adding a timeline to the manipulation it is possible to get a pulsing, alternating, or sporadic repeated change effect of colour, rotation, or even size. Here, the emissive properties of the stimulus in the environment can be altered, to be used as an indicator to the participant that they have pressed an invalid button on their controller. Before adding the timeline (outlined in *Section 2.4.3*) the dynamic material instance needs to be created: find the desired material in the `Content Browser`, double click and add an open parameter to the property that is going to be manipulated. Here we will add `ScalarParameter` to `Emissive Color` and rename to "EmissiveOutput": this can be seen in *Figure 2.14*, but any number of adjustments can be made to the material properties by following these steps. Return to the scene and in the `Content Browser`, right-click on the chosen material and create a material instance. Then, in the `Level Blueprint`, add the `Create Dynamic Material Instance` node and for the

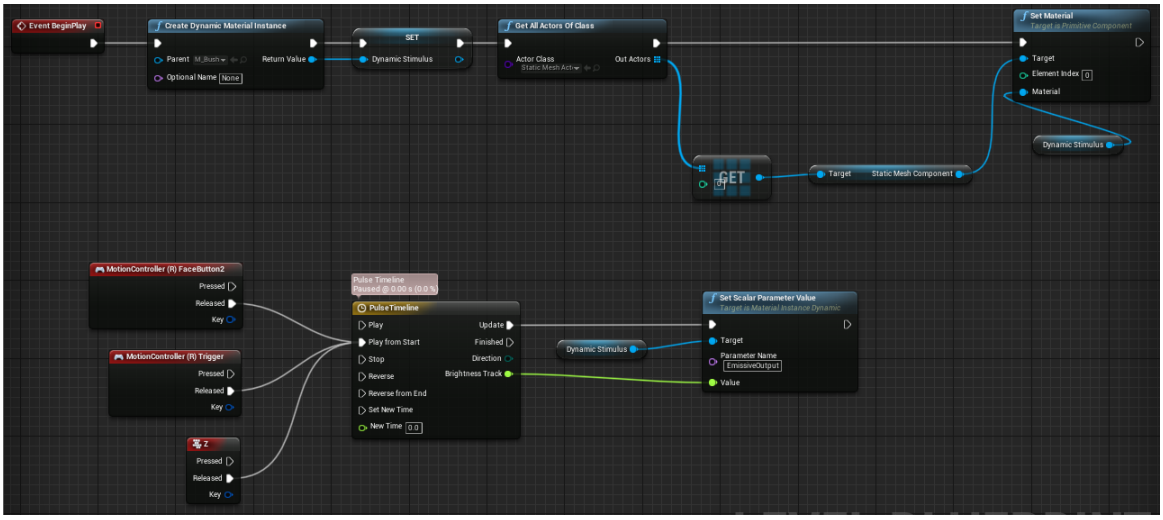


Figure 2.13: Two flows needed in order to create a dynamic material. The material instance is created at the start of game play, and on button presses the material is set to adjust based on a timeline.

`Parent` input, select the original material. In this example, we can select `M_Bush`. Next, on the `Return Value` output, right-click and promote to a variable, we can call this "DynamicStimulus", connect this to a new node `Get All Actors of Class` and connect that to a `Set Material` node. In `Actor Class` it is also necessary to select the appropriate class of the target object that will be changing: in this example it should be 'StaticMeshActor'. For the `Out Actors` output on the `Get All Actors of Class` node, connect the `Get (a copy)` node, and then connect this to the `Target` on the `Set Material` Node: *Figure 2.13* presents this. One the dynamic material has been created, an input or output event (*Section 2.4.2*), or a timeline (*Section 2.4.3*) can be attached.

2.4 Timings

2.4.1 Using Variables to Keep Track

Variables are properties that hold a value or reference an Object or Actor in the world; commonly used variable types include Boolean, Integer, Float, String, Object. Creating a variable to keep track of the number of trials completed, or whether or not a condition has been completed, are useful ways to incorporate variables to ensure smooth running

of an experiment.

Instruction: create trial counter

Create a new variable in the `Level Blueprint` to keep track of the current trial number.

```
Toolbar > Blueprints > Open Level Blueprint > My Blueprint >
Variables > + Variable > *rename ("TrialCounter")* > Details >
Variable Type > *select Integer*
```

Repeat for another variable with pre-defined total number of trials.

```
My Blueprint > Variables > + Variable > *rename ("TrialsTotal")* >
Details > Variable Type > *select 'Integer'* > Toolbar > Compile >
Details > Default > Total Trials > *adjust to "5"*
```

Start a flow to come back to, to keep track of when a trial is completed.

```
My Blueprint > Variables > *drag and drop 'TrialCounter' to
blueprint* > Set TrialCounter > My Blueprint > Variables > *drag
and drop 'TrialCounter' to blueprint* > Get TrialCounter > *drag
'Trial Counter output', release mouse click* > *search '+'* >
integer + integer > *link '+ output' to 'Set TrialCounter'*
```

Variables are very useful at keeping track of whether a trial has been completed and the experiment can therefore move onto the next trial, or not. They can be Boolean (true/false), integer (where the value is pre-defined or altered during the experiment), or other formats, for example strings of text and arrays, which are particularly useful when recording data (*Section 2.7.3*). *Figure 2.15* is an appropriate flow to create here: by creating a Boolean variable ("IsTrialFinished") in the `Level Blueprint`, the default value can be set (to 'False'/unticked), and when a certain condition is met (participant presses a button as a response) the variable value can be automatically changed (to be 'True'/ticked). Instances like this are where the `Branch` node comes in useful; think of this node as an *if statement* in coding. We can add this to the flow to only move on to the next trial when the variable is equal to one, indicating the trial has been completed. The same logic can be applied to checking whether conditions have been completed.

Instruction: track completed trials

Create a flow similar to *Figure 2.16*, adding the `Equal (integer)`, `Branch`, and `Quit Game` nodes.

```
Level Blueprint > My Blueprint > *drag TrialsTotal to Event Graph*
  > Get TrialsTotal > *drag output of TrialsTotal, release mouse
 click* > *search "=="* > Equal (Integer) > My Blueprint > *drag
   TrialCounter to Event Graph* > Get TrialCounter > *connect
            'TrialCounter' to 'Equal (Integer) B'*
```

Ensure that the application consistently checks whether the correct number of trials have been completed. This will be extended later on to end the experiment when the condition is met.

```
 Event Graph > *right-click empty space, search and select 'Event
Tick' > *drag 'Event Tick Exec', release mouse click* > *search and
 select 'Branch'* > *link '== Return Value' to 'Branch Condition'
                          Boolean*
```

2.4.2 Events and Trials

An important node to include is `Event BeginPlay`: it is not necessary to have a series of observable actions stemming from this node, however its function in the blueprint is to start any features as soon as the game/experiment begins. A visible use of this node would be to connect it to the dynamic material flow detailed in *Section 2.3.4* so that the material instance is automatically set up and ready to go as soon as the experiment begins.

There are other occasions where an event is desired to occur when there is a specific onset criterion to be met. A simple example of this can be seen in *Figure 2.12*: when the key 'P' on the keyboard is pressed the material of the back wall changes, and the following events also occur due to the keyboard press initiation. It is often useful to use a keyboard press from the experimenter as a way to start the experiment, like in this example, as well as in the tutorial instructions.

In certain instances, an event would be desired to occur over and over again: the `Event Tick` node calls on the connected nodes on every frame of gameplay. As can

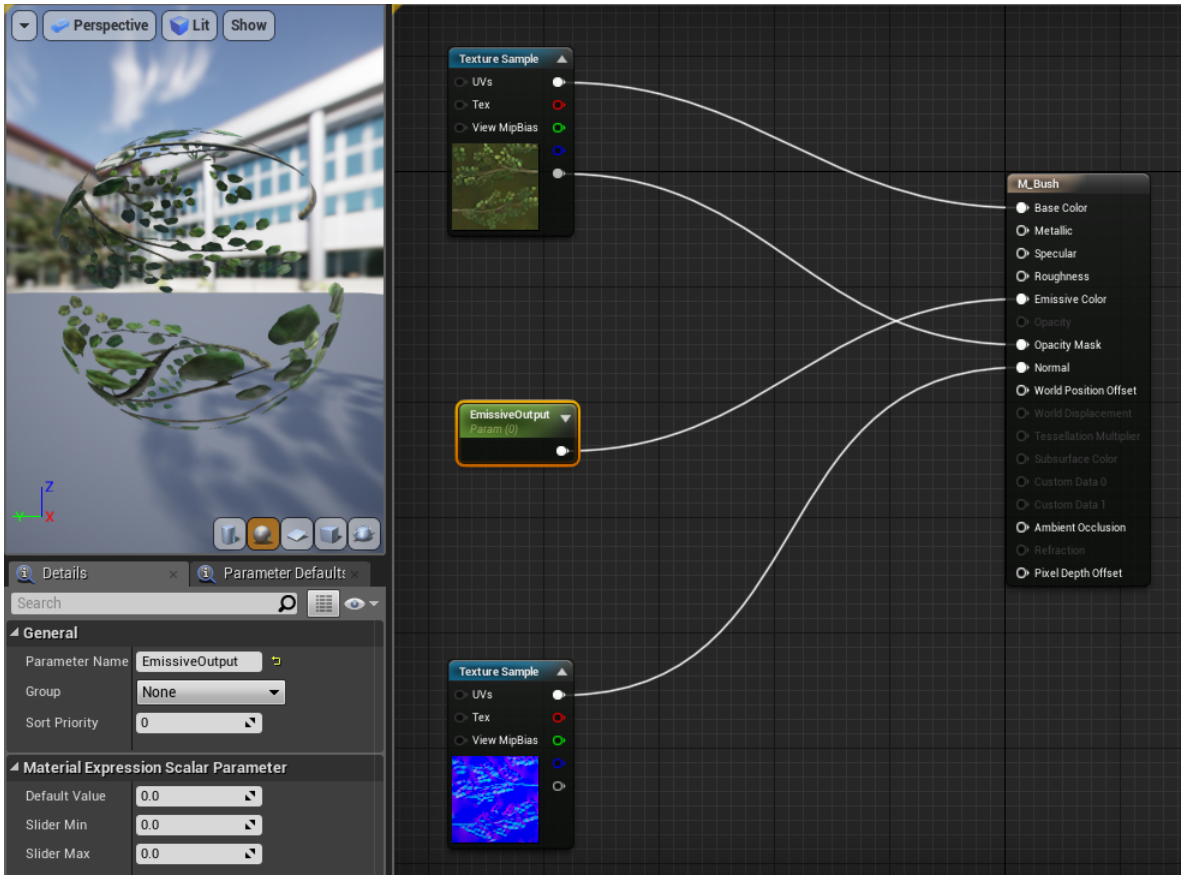


Figure 2.14: The material properties of an object places in the environment, where the Emissive Output is set to be adjusted in a dynamic material instance.

be seen in *Figure 2.16* the `Event Tick` is continuously checking whether the variable labelled `Trial Counter` is equal to the variable `Total Trials`, when the two variable integers match, the flow follows onto `Branch` to then save the data and end the game. This is a quick way to ensure automatic ending of the experiment once a predefined number of trials in `Total Trials` is set.

2.4.3 Creating Timelines

The use of timelines is helpful where certain occurrences are desired to happen at set intervals. Examples of this might be a light to flicker on and off every x seconds, or a stimulus to tilt x degrees in one direction before tilting in the opposite direction. Because the nature of a timeline is specific, the node can only be obtained within a base/level blueprint (as well as pawns and actors) as they need to be linked to the engine tick

which widget blueprints, UMG, and objects are not.

A scenario which demonstrates this is the experiment by Yuan and Steed (2010) where after 10 minutes of playing engaging (distraction) games in VR, a lamp would appear to fall onto the avatar's hand. Here, the amount of time before this event is needed to occur is known, which is why it is a great example of where a Timeline could be used.

If a pulse of some kind is desired, such as sinusoidal change in luminance of a stimulus (*Section 2.3.4*), the following steps would be appropriate to code; however, slight adjustments to the flow can be made if movements, rotations, etc are the desired manipulations. First, insert the `Add Timeline` node, it is then possible to rename ("PulseTimeline"), double click this to make the necessary manipulations to the timeline. Clicking the `Add Float Track` on the upper left corner of the timeline will add the timeline tracker (the same applies if adding a Vector or Colour track). This can be named "BrightnessTrack", and will appear as the output on the node itself. The editable value in `Length` is the total length of the curve/ the time taken before this sequence repeats. By right-clicking anywhere on the timeline track and selecting 'Add key', we can then adjust the 'Value' and 'Time' (that this happens in the timeline) of this by inputting numbers in the appropriate boxes at the top of the track. To create a pulse, we can add keys at the beginning of the timeline, the middle, and at the end. By setting the value of the first and last key to zero, and the middle key to three this will create a pulsing effect, which can be smoothed out if each key is right-clicked and 'Key Interpolation - Auto' is selected. This timeline is shown in *Figure 2.17*.

The timeline gets added as a `Blueprint Component`, so it is possible to get references to this elsewhere in the blueprint. If this pulse is wanted to repeat each time a trigger occurs (*Section 2.4.2*), the trigger should be linked to the node input `Play from Start` as shown in *Figure 2.13*. Alternatively, if one trigger is intended to cause a continuous loop, the `Loop` icon in the timeline should be selected. You need to save and compile

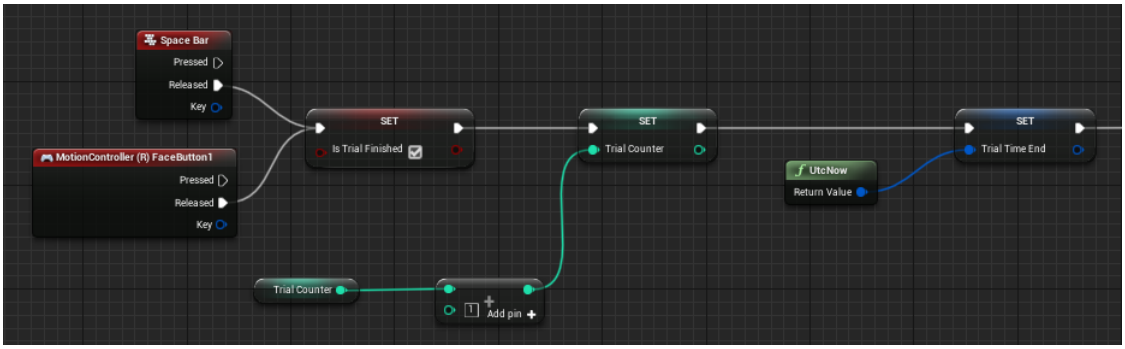


Figure 2.15: When a specific button is pressed, indicating the participant has finished the trial, the variable to keep track of the number of trials is adjusted and the current time is recorded as another variable.

before attempting to view any changes made, otherwise it is likely that they either will not show up, or an error message will pop up.

A reference to the target and the `Set Scalar Parameter Value` can then be linked to this timeline node. The `Parameter Name` should be edited to whatever feature of the target object is desired to be changing, in *Section 2.3.4*, 'EmissiveOutput' is chosen as the parameter name, so this should be spelt exactly the same here. In *Figure 2.13* the emissive strength on the *Stimulus* is to change from zero to one when 'Z' on the keyboard is pressed (for debugging purposes) and other controller buttons which are not in use throughout the rest of the experiment. The connection is made from the `Update` output here, however if another occurrence is desired to happen once this has finished, the flow can be continued from the `Finished` output from the timeline. This can be thought of as a *fire and forget tick* as nothing else will occur through the use of this timeline. Now, we have created a debug to allow the participant to know they have pressed an incorrect response button during the experiment.

2.4.4 Keeping Track of Time

The `Timeline` is a helpful aspect when programming a series of events to happen over a set period of time, however in many experimental procedures it is crucial to keep track of the time passing during the experiment, than the time between events. There are a few different ways to set up a time, one is to use the timer handle and counter, although

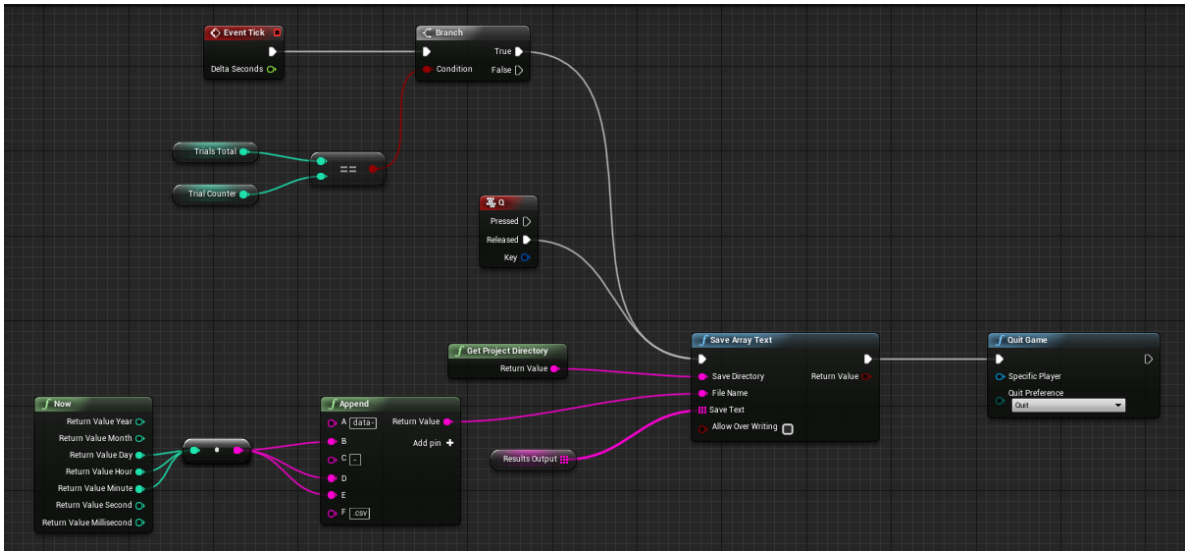


Figure 2.16: Example end of level set-up, where the application will save the results and close down once all trials and conditions have been completed, or Q is pressed.

this is not a very accurate way to keep track as it largely depends on the current refresh rate.

An alternative, for the most accurate tracker which is unaffected by any refresh or lag of the equipment, is to use the `Now` node to get the exact time the experiment begins (as precise as milliseconds) and then again at the end of the experiment. It is common for journal articles to publish the total time taken for a single participant to complete the experiment (Kytö et al., 2013; Schwind et al., 2019; Yan et al., 2018). In some experiments the time can extend to multiple hours, but the range varies greatly depending on the paradigm. In a by Asher and Hibbard (2018), participants were required to return to the laboratory a minimum of two times to complete the experiment. This was the case because there was a total of 18 trial blocks, with each block containing 120 trials: too many to complete in one sitting. The overall time taken for participants, across the sessions, were recorded and presented in the write-up.

Reaction times are objective rather than subjective indicators of cognitive performance (MacLeod, 1999). In a meta-analysis of 114 studies by Suchotzki et al. (2017), reaction times were obtained from 3307 participants in order to establish whether lying was more cognitively demanding than telling the truth. It was found that verbally ly-

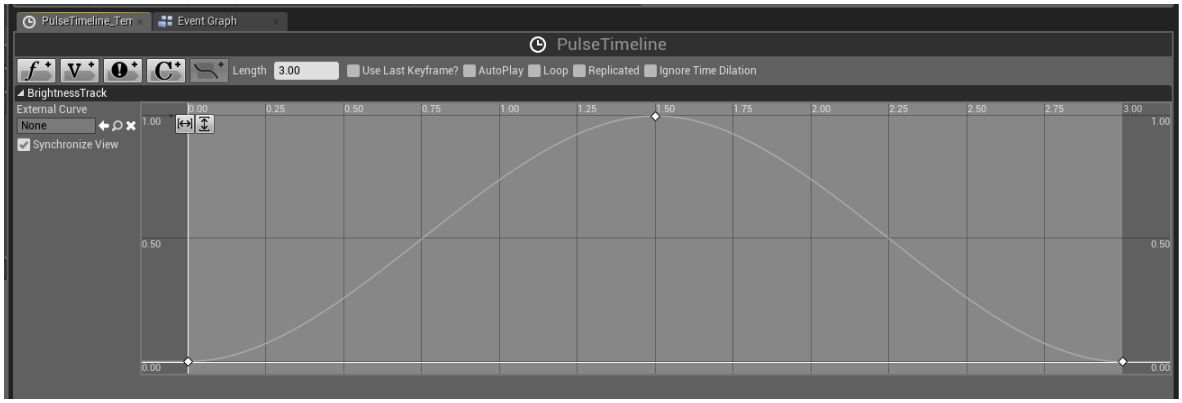


Figure 2.17: Timeline of 3 seconds, where the brightness steadily increases and decreases.

ing takes significantly longer to come up with a response than when responding with a truthful response, even in scenarios where the participants were told to conceal their lie as much as possible.

Setting up an experiment to record the time taken between any two events can be set up in the same way as keeping track of the overall time taken to complete the task. The the `Now` time recorder in this instance should be positioned first at the onset of any stimuli (the start of the trial loop), and then again at the end of the trial, or when the variable `IsTrialFinished` is set to *True*. When it then comes to data analysis it will be possible to measure the time between these two times: the reaction time for that trial. It is important to ensure these are saved to the data output, which is explained in *Section 2.7.3*. After inserting the `Now` node, right-clicking on 'Return Value' and selecting 'Split Struct Pin' makes it possible to select exactly which time values are to be recorded. Alternatively, you can not split the pin and save the date and time as a whole. It is important to right-click on this node and promote it to a variable, so that this information can then be stored and can be written into the data output (*Section 2.7.3*).

2.4.5 Setting Time

It is often useful to program a set time into a flow: in *Figure 2.12* the `Delay` node is used to create a 2 second delay after the material of the back wall is altered, before moving

on to the next features of the experiment. The *Duration* input float can be adjusted the desired number of seconds, or a variable can be hooked up here with information predefined elsewhere, or even a random number (*Section 2.8*).

There are many nodes that can be used to create a random number, for instance the `Random Integer` can be used to obtain a number between zero and a set maximum (the maximum can also be linked to a variable defined elsewhere). This is useful for staggering and randomising the delay time between trials. The `Random Float` is very similar, the difference being that here a float is used, which is a number that includes decimal places, rather than a whole number (integer). The use of the random number will directly affect which node is to be used, for example if the rotation of a stimulus is required to be randomised, the `Random Rotator` is necessary as it will output all X/Y/Z (roll/pitch/yaw) values. Additionally, the `Random Bool` can be used to randomly decide whether the linked nodes are to be 'True' or 'False'.

2.5 Conditions

2.5.1 Sequential

Depending on the paradigm being tested, it is sometimes necessary to test participants in a specific order of circumstances (Webb et al., 2018), or present a set of conditions in a strict order. It is simple to set up UE4 to present conditions in a non-randomised order, once the order has been decided on and conditions created. In *Figure 2.18* the conditions are collapsed so that the overall procedure of the method is more legible (more information in *Section 2.8*) with each condition-node containing the more specific parameters such as number of repetitions, adjusting stimuli specifics, etc. The `MultiGate` node should be used to set the order of the conditions, from first (top) to last (bottom). This node also has the option to loop over again by selecting the 'Loop' Boolean if repetitions of the outputs are necessary.

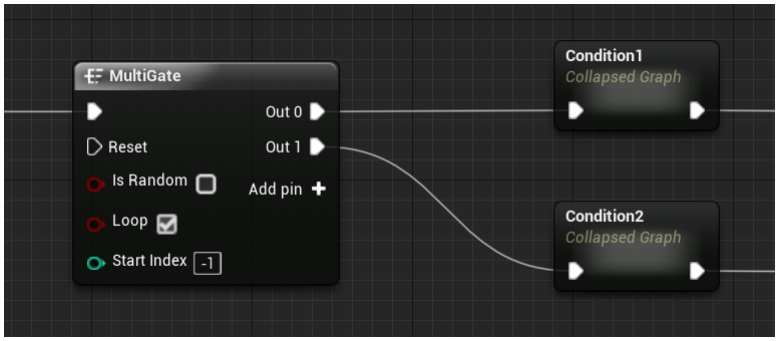


Figure 2.18: MultiGate node used to loop through two conditions in order; the conditions have been collapsed.

2.5.2 Order Randomisation

The `Is Random` Boolean on `MultiGate` can be toggled on or off, which allows for another sequential presentation order of conditions if toggled off. If toggled on, the order of each condition linked to the output is randomly picked. In MATLAB it is necessary to plant a randomisation seed, to ensure that the first round of random order is not the same each time the programme starts; there is no documentation to suggest this is necessary in UE4, and repeated pilots have provided as evidence that the first round is not the same every time.

2.5.3 Counterbalancing

Counterbalancing provides an equal proportion of participants being exposed to each possible condition first. Complete counterbalancing (presenting all possible orders to equal numbers of participants) is possible if there is a small number of conditions. If not, then techniques such as a Latin square design can be used to remove some of the more obvious ordering effecs, such as equating the number of times each condition is presented first. These techniques are useful for reducing order effects. Yan et al. (2018) explored the feasibility of users performing eyes-free target acquisition in near-space. Counterbalancing between participants being seated or stood up in two sessions was used, it was found that there was no effect of posture, other than the subjective comments of having more stability when seated. Without counterbalancing, it would

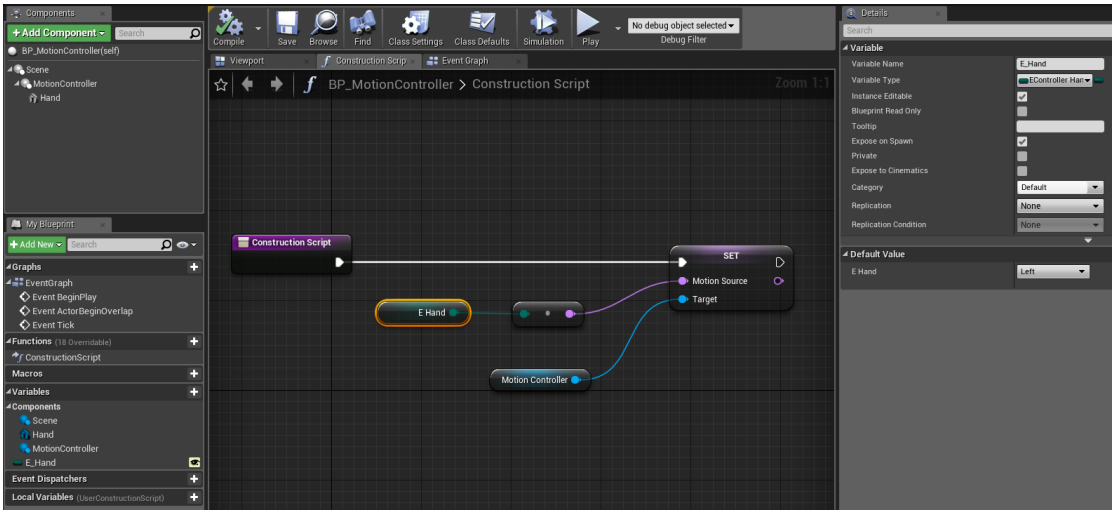


Figure 2.19: View of the blueprint when setting up controllers for a First Person C++ project.

not have been possible to dissociated posture effects from any possible ordering effects, such as an improvement in performance in the second condition due to training on the task.

Much like order randomisation, this can also be achieved through the `MultiGate` node if two output sequences are connected and the `Is Random` Boolean is selected. This will not guarantee an equal proportion of participants will experience one condition first, so a more accurate substitute would be to utilise the experimenter's ability to interact and set up the experiment. As describes in *Section 2.4.2* we can start specific occurrences after button presses so a simple fix would be to use one keyboard button as a way to instigate the conditions from one starting point, and another button press to start from the alternative.

2.6 Participant Interaction

2.6.1 Controller Avatars

The mode which participants use to interact during an experiment will usually be dependent on the task itself. Many experiments use keyboard key presses, mouse clicks, or controller button presses as the interaction method for participant responses; touch-

screens have also been found to be better than many other interfaces for certain methodologies (Brasel & Gips, 2014). A study by Werkhoven and Groen (1998) found that hand tracking, compared to a 3D-space mouse input device in VR, yielded better performance. This input system would not be appropriate for more vicarious set-ups, so the type of paradigm is needed to be considered before determining a controller system.

Motion controllers refer to the gaming controllers associated with the headset. Participant responses sometimes require pulling of the trigger (Sarupuri et al., 2017; Smith et al., 2015) or physical movement, tracked by the controller (Geuss et al., 2015), in the environment to complete a trial.

The blueprint for the `MotionControllerPawn` has the controller information set-up and ready to go, and the steps taken in *Section 2.2.1* will automatically allow for the tracking of the controllers in the virtual environment. However, the following can be used to create the equivalent, if using a First Person C++ project (as described in the instruction tutorials) as opposed to VR Blueprint then this can be incorporated via the following instructions. First, a blueprint for the virtual representation of the participant (avatar) needs to be created:

Content Browser > ForExperiment > *right-click on empty space* > Create Basic Asset > Blueprint Class > Pawn > *rename ("BP_Participant")* > *double click to open this new blueprint* > *option to change scene root: *Section 2.8** > Components > Add Component > New Blueprint Script Component > Scene Component > Next > *rename ("VRorigin")* > Create Blueprint Class > *return to 'BP_Participant* > Components > Add Component > Camera > Toolbar > Save > Compile

Next, a representation of the hands needs to be created:

Content Browser > *right-click on empty space* > Create Basic Asset > Blueprint Class > Actor > *rename ("BP_MotionController")* > *double click to open this new blueprint*

Components > *option to change scene root *(described in Section 2.8)* > Add Component > MotionController > Add Component > Sphere > *rename (Hand)* > *drag and drop 'Hand' onto 'MotionController'* > Details > Transform > Scale > *adjust sizing along XYZ axes to '0.2'* > My Blueprint > Variables > + Variable > *rename (E_Hand)* > Details > Variable Type > *search and select 'EController Hand'* > Toolbar > Save > Compile

Next, select the `Construction Script` for this blueprint. Format the controllers as spheres, as shown in *Figure 2.19*.

```
Components > *drag MotionController to script* > *drag the output* >
*search and select 'Set Motion Source'* > My Blueprint > Components >
*drag E_Hand to script* > Get E_Hand > Details > *select 'Instance
Editable'* > *select 'Expose on Spawn'* > *link E_Hand to 'Motion Source
Name'* > *connect 'Construction Script' to 'Set Exec'* > Toolbar > Save
> Compile
```

Next, return to the `BP_Participant` blueprint, in the `Event Graph`.

```
*drag from 'Event Begin Play', release mouse click* > *search and select
'Branch'* > *right-click 'Branch Condition'* > Promote to variable >
*rename variable ("SpawnControllers")* > *drag 'Branch True Exec',
release mouse click* > *search and select 'Sequence'
```

The following should be repeated twice: once for the left controller, pulling 'Sequence Then 0' and naming nodes as 'L'/'Left'; once for the right controller, pulling 'Sequence Then 1' and naming nodes as 'R'/'Right'.

```
*drag 'Sequence Then 0', release mouse click* > *search and select
'Spawn Actor from Class'* > Class *select 'BP _MotionController'* > E
Hand *select 'Left'* > *right-click 'Spawn Transformation* > Split
Struct Pin > *rick-click 'Return Value'* > Promote to Variable > *rename
("BP_MotionControllerL")* > *drag 'Set Exec', release mouse click* >
*search and select 'AttachToComponent'* > *drag 'Set output' to
'AttachToComponent Target'* > Components > *drag 'Camera' to Event
Graph, release mouse click* > *link 'Camera' to 'AttachToComponent
Parent'* > *adjust 'Location Rule' and 'Rotation Rule' to 'Snap to
Target'*
```

From the 'Then 1' Exec of the `Sequence` node this can be repeated for the right hand, adjusting variable names accordingly (swap L for R and Right in E Hand). This is depicted in *Figure 2.20*.

The sphere representing the controllers used in the tutorial can be replaced with anything: avatars for hands can be downloaded from external sources (*Section 2.2.2*) and applied, or the scene root can be altered so that there is no visible avatar for the participant (*Section 2.8*).

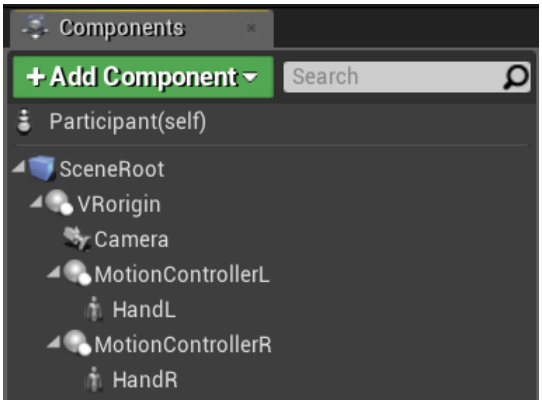


Figure 2.20: The Component tab within the Participant blueprint depicting the camera, motion controllers, and visible hand meshes.

Some paradigms need to dehumanise the avatar, or reduce agency in a virtual environment. This can be achieved by positioning a virtual object in place of the hand avatar (Grubert et al., 2018) or introducing a lag to the avatar (Rohde et al., 2014; Van Dam & Stephens, 2018). A study by Werkhoven and Groen (1998) found that in an immersive display, when the hand avatar was correctly aligned with the participant's hand/controller, performance was better than when it was misaligned. Conversely, it has been found that an offset of the virtual hand doesn't hinder performance on visuo-motor control task (Burns et al., 2005; Groen & Werkhoven, 1998). Not only this, but Yuan and Steed (2010) investigated whether the *Rubber Hand Illusion* could be replicated in VR, by manipulating the distortion (centimetre offset) on the avatar hand associated with the participants' hand, while also measuring Galvanic Skin Response when a prop collided with their avatar. All of these manipulations are easily achieved in VR, and through the use of UE4, however the use of external equipment to measure biometrics will not be described here but set-up instructions should be found in the apparatus' user manuals.

2.6.2 Virtual Responses

It may be desirable for participants to make their responses within the virtual environment. Examples of this might be to adjust an icon on a virtual Likert scale, as this gives a larger range of possible outcomes than simply pressing trigger or face buttons

on a controller; or once environment collision (*Section 2.6.3*) has been set up, using the motion controller to hit a virtual response in the environment. In a study by Kytö et al. (2013), after the trials of the experiment participants were asked to rate on a Likert scale how much a visual component aided in their decision-making throughout the task. This could have been made more efficient by including the scale as a pop-up within the task after each trial/block of trials, to avoid the risk of participants forgetting features of the experiment which might have aided their Likert answers.

Before creating any virtual responses, the exact key or button presses need to be identified, and what the desired result will be. When selecting any of the `Keyboard Events` or `Gamepad Events` in the `Event Graph`, outputs can be linked from when the button is pressed or released. Often, it is good measure to link the output with the 'Released' pin, so that whatever is linked in the flow does not keep repeating if the button is held down. If an adjustment to an asset in the environment is to occur once a button is pressed, a reference to the asset needs to be included in the flow: by selecting the object in the `Level Editor`, moving to the `Level Blueprint` and then right-clicking on an empty space in the `Event Graph`, a reference can there be created. Adjustments to assets in the environment can be linked to button pressed: the `SetActorRotation` node can be used to rotate an object; the `SetActorLocation` can be used to move one; for both types of adjustment and for sizing, the `SetActorTransform` can be used. If the `SetActorTransform` is used, the 'New Transform' pin should be split (right-click on 'New Transform' and selecting 'Split Struct Pin'). This is done so that each type of transformation can be edited separately as floats, if this is not done then whatever number is inserted will adjust the specified XYZ values together.

Often it is useful to make relative adjustments to assets, rather than absolute: by using the `GetActorRotation` (or `Location`, or `Transform`) the current rotation of the stimulus is identified. Splitting the the `Return Value` will then display the XYZ co-ordinates, and some sort of numeracy node can then be added in to calculate the extent of change to be made. In *Figure 2.10* this is used, along with a transformation to the location of an asset: the `Random Float in Range` node is used to create an adjustment

to the location of the `Target` along the X axis, while retrieving the current Y and Z locations to ensure there is no change along these axes.

Instruction: format participant response

Create flow to format movement of the stimulus.

```
*select ‘Stimulus’* > Toolbar > Blueprints > Open Level Blueprint >
  Event Graph > *right-click* > Create a Reference to Stimulus >
  *drag output from ‘Stimulus Reference’, release mouse click* >
      *search and select ‘SetActorRotation’* > *right-click
‘SetActorRotation New Rotation’* > Split Struct Pin > *drag output
  from ‘Stimulus Reference’, release mouse click* > *search and
select ‘GetActorRotation’* > *right-click ‘GetActorRotation Return
  Value’* > Split Struct Pin > *drag output from ‘Return Value Z
(Yaw), release mouse click* > *search "+"* > float + float > *type
      "10" into the ‘+ B float’* > *link ‘+ Return Value’ to
‘SetActorRotation New Rotation Z (Yaw) on ‘SetActorRotation’ node*
       > *link ‘GetActorRotation Return Value X (Roll)’ to
‘SetActorRotation New Rotation X (Roll)’* > *link ‘GetActorRotation
    Return Value Y (Pitch)’ to ‘SetActorRotation New Rotation Y
                          (Pitch)’*
```

Add controller button presses to rotate the stimulus, as depicted in *Figure 2.21*.

```
*right-click on empty space of blueprint* > *search "keyboard a"* >
 A > *drag ‘A Released’ to ‘SetActorRotation Exec’* > *right-click
    on empty space of blueprint* > *search "motion trigger"* >
MotionController (R) Trigger > *drag ‘MotionController (R) Trigger
              Released’ to ‘SetActorRotation Exec’*
```

2.6.3 Environment Collision

Collision is important when creating virtual environments which intend to present a scenario with object interactions. A study by Holmboe et al. (2019) explored inhibitory control by presenting various stimuli on a tablet: participants were to press the screen when and where a particular stimulus appeared. In this particular experiment, a large proportion of the participants were under the age of ten, so it would have been unsuitable to use VR, however this type of method could of course be adapted to allow participants to reach out and touch stimuli when they appear in a virtual environment, with the time

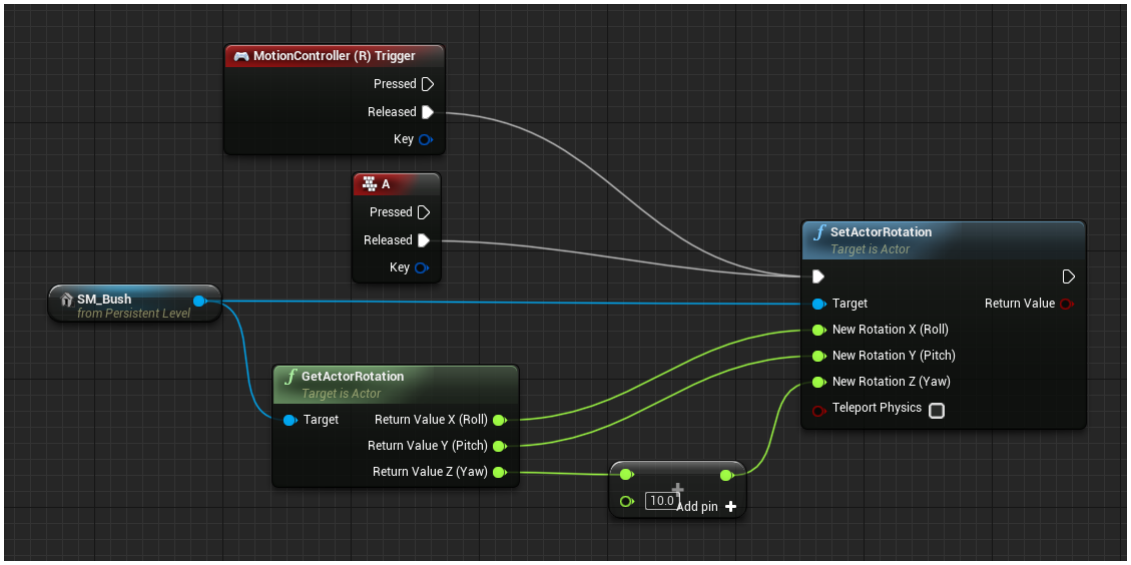


Figure 2.21: Flow where the rotation of the Stimulus is retrieved and manipulated by button presses.

of controller/object collision saved in each trial.

Coding an experiment where objects collide with one another is especially easy when stimuli are obtained via the marketplace or are simple structures from the Starter Content, as these will usually be accompanied by collision built-in. However, in instances where this is not the case, adding collision to an object can be as simple as double-clicking on the static mesh of said object, locating the `Collision` option at the top of the window, then selecting the appropriate type of collision. A collision boundary is an invisible area which surrounds the object; when two (or more) virtual objects which also have a collision boundary come into contact with one another, they will be able to collide but not pass through one another. This is necessary when a player avatar is walking on a floor plan, both the avatar and floor need collision so that the player does not fall through the floor. `Simple Collision` will cause the collision boundary of the selected shape to be a basic shape surrounding the overall, general shape of the object, whereas `Complex Collision` will create a more refined and accurate boundary for the collision to occur around. As can be seen in *Figure 2.22*, `Complex Collision` is selected. Having this set up on two objects, such as a floor and a falling object, can mean the difference between objects staying in the environment during the experiment

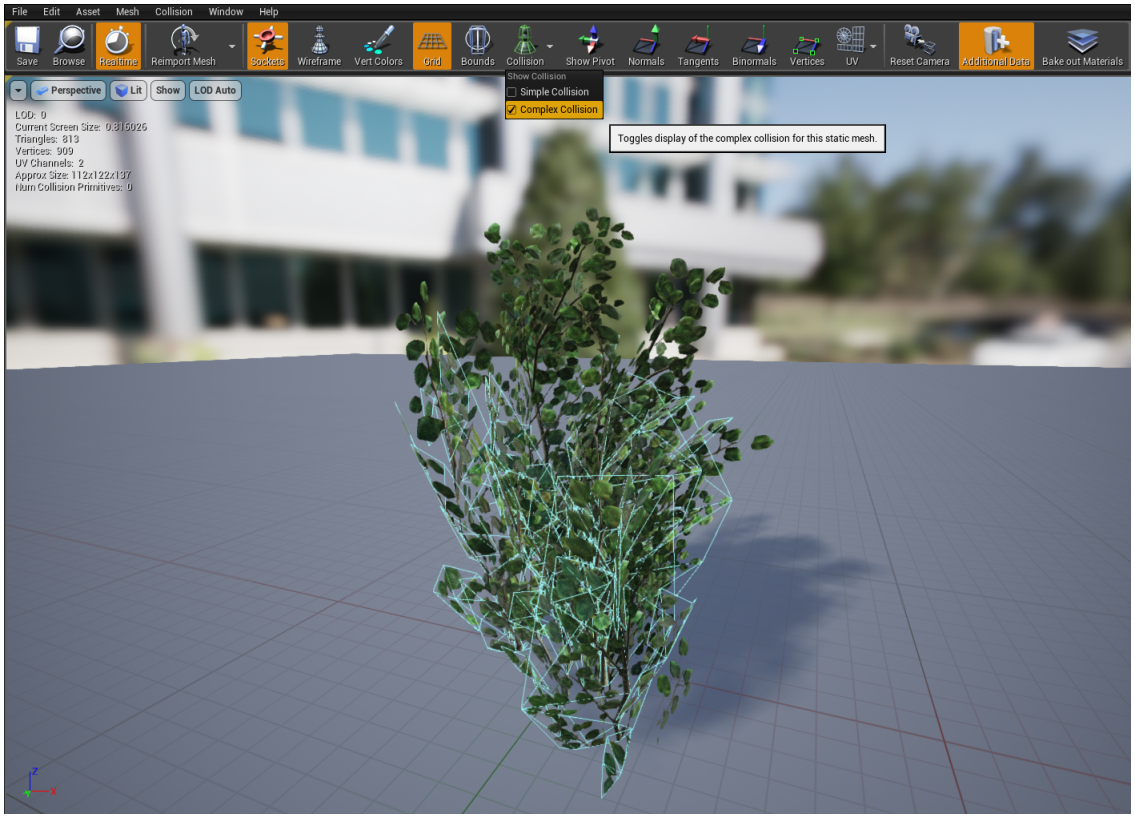


Figure 2.22: The static mesh of the bush is shown, with complex collision added. The area of collision is represented by the cyan lines.

and them unintentionally moving out of sight through the floor.

Having this added component of natural physics can induce a sense of immersion: Slater and Steed (2000) found that participants who were able to interact with their virtual environment have a higher sense of presence than when merely touching a response button. This would be an ideal application for collision within the environment.

2.7 Data Collection

2.7.1 Obtaining Raw Data

The reason for creating an experiment in UE is to collect data for analysis: therefore it is essential that the right information is collected. Previous steps have explained how to transform information into variables which can later be accessed and/or altered, this is also a good way to temporarily store information which is desired for saving when the

experiment ends.

- *Collecting stimulus variation:* the experiment can be set in such a way as to present a random stimulus with a predetermined set of variations (*Section 2.4.5*) or at specific XYZ coordinates. Within a trial, you can create a new variable and adjust its type to a `Vector`, then pull it into the blueprint and `Set` it by connecting the `GetActorLocation` node (with a reference to the stimulus) to the input. Now in each trial this variable will be adjusted to these coordinates.

- *Collecting observer responses:* once the controllers have been set up, it is simple to get the response from each trial by setting the response to be a variable, much like the preparation for collecting the stimulus variation. Instead of getting an actor's location you can set a variable, from participant controller button presses for example. In the `Level Blueprint`, link the controller press to the `Set Variable` node for your new variable. Here you can set the variable to to be a Boolean (trial completed 'True'/'False') or Integer ('0'/'1'), to keep track which button has been pressed by creating a reference key and allocating different numbers to different buttons.

- *Collecting additional metrics:* additional data such as reaction time is easily added to the trial recordings, as described in *Section 2.4.4*. On the other hand, if there are multiple conditions being tested then variables can be adjusted to keep track then a key should be created, with the numbers that are associated with individual condition variations In doing this, the variable number can be tracked back to which stimulus was presented when analysing the data. Within the blueprint, a variable can be coded to be set to each of these numbers in the key, when each differing condition is presented. This variable can be added as an additional output in the saved data (*Section 2.7.3*) and then when analysing the data output, the exact stimulus which was presented in each trial can be deciphered.

Instruction: create an array for the raw data

Data collected from each trial will be separated by commas and aligned by row: first column will be trial number, second column the stimulus rotation.

```
*ensure ‘Stimulus’ is selected* > Toolbar > Blueprints > Open Level
   Blueprint > *right-click on empty space* > *search and select
‘GetActorRotation’* > *right-click ‘GetActorRotation Return Value’*
> Split Struct Pin > *right-click empty space* > *search and select
 ‘Append’* > *type "," into ‘Append B’* > *right-click ‘Append’* >
 Add pin > *link ‘GetActorRotation Return Value Z (Yaw)’ to ‘Append
                                C’*

My Blueprint > Variables > *drag ‘TrialCounter’ to the Event Graph*
 > Get TrialCounter > *link ‘TrialCounter output’ to ‘Append A’* >
 My Blueprint > Variables > + Variable > *rename ("ResultsOutput")*
 > Details > Variable > Variable Type > *select ‘String’* > *click
 the pink dash next to the Variable Type* > Array > *drag and drop
  ‘ResultsOutput’ to the Event Graph* > Get ResultsOutput > *drag
  from ‘ResultsOutput output’, release mouse click* > *search and
   select ‘Add’* > *link ‘Append Return Value’ to ‘Add New Item’*
```

2.7.2 Thresholds

Perceptual thresholds, specifically in vision science, are typically retesting while varying one dimension of the stimulus: Owsley and Sloane (1987) tested contrast sensitivity thresholds on real-world targets, to find a link between spatial contrast and detection of stimulus type. They tested five spatial frequencies on objects which were deemed to be typical targets in everyday life. When it came to data analysis, the experimenters had records of the stimulus, contrast condition, and participant response, so were able to perform appropriate regressions to estimate psychometric functions, describing how responses varied with contrast, for each spatial frequency.

In order to calculate the PSE and JND thresholds, the data which can be collected from an experiment should be: (1) the stimulus variation level; (2) the response; (3) optional additional metric such as galvanic skin response (Ayata et al., 2017; Meehan et al., 2003) or response time (Ghisletta et al., 2018; Roitman & Shadlen, 2002).

Calculating thresholds tends to be a procedure which happens in the analyses stage of experimentation. Therefore, no specific calculations or coding needs to take place during the experiment, however effective storage of all stimuli, responses, and trial count are important and are discussed in *Section 2.7.1* and *Section 2.7.3*.

2.7.3 Saving Data to File

As UE is not typically built to save data, a custom node needs to be programmed. This requires installation of Visual Studio to write some custom C++ coding, however it is only a matter of copying the text in *Figure 2.23* which shows the two Notepad++ file contents to create the node. By adding this code to the level, a new node named `OutputResults` can be inserted into the UE project multiple times, however it is typically only needed once to save data when the experiment is completed or exited for whatever reason (participant may want to take a break, or need to come back to complete at a later date).

Instruction: create a data saving node, part 1

A new class for the node should be created.

```
File > New C++ Class > *select 'Show All Classes'* >
BlueprintFunctionLibrary > Next > *name function ("OutputResults")*
> *select 'Public class'* > Create Class
```

Visual Studio should automatically open (if not, the file can be found in the File Explorer for the PC: \Documents\UnrealProjects\experiment1\Source\experiment1\Public) and the following will be written in the 'OutputResults.h' tab, approximately on line 17 and saved. Some should be written on new lines, *Figure 2.23 (a)* can be used as a reference.

```
UFUNCTION(BlueprintCallable, Category = "Custom", meta = (Keywords
= "Save"))
static bool SaveArrayText(FString SaveDirectory, FString FileName,
TArray<FString> SaveText, bool AllowOverWriting);
```

Ensure this is saved before moving on.

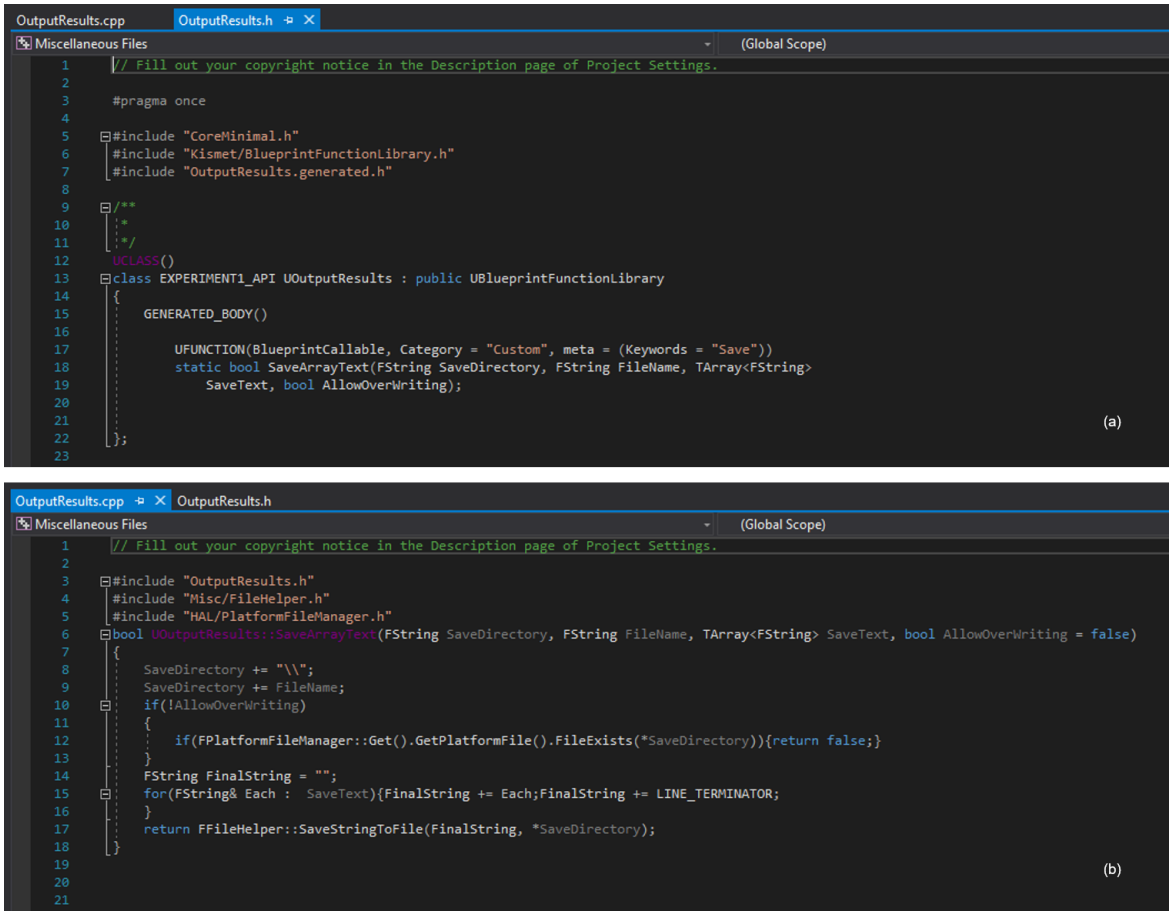


Figure 2.23: Screenshot example of the format coding described in *Section 2.7.3*, with the OutputResults.h in (a) and OutputResults.cpp in (b).

In line 18 of the 'OutputResults.h' script, the section 'SaveArrayText' is used so that the node does not save one single cell: an array will allow for more rows/columns to be included. The items within the parenthesise are to be the inputs for the new node. When creating a new document to save, it is desirable to be able to have control over where the file is saved to ('SaveDirectory'), the name ('FileName'), and whether or not a new file should overwrite an existing file should they both have the same name ('AllowOverWriting'). Ensure you save this before moving on.

Here, the node has been set so that if a new file is to be created with a duplicate name in the same directory, it will not overwrite the original. This is important if each file will contain responses for a single participant, and it has been coded to name the file a generic name (for example, "Results").

Instruction: create a data saving node, part 2

The following is to be written in the 'OutputResults.cpp' (which can be found at: ...\Documents\UnrealProjects\experiment1\Source\experiment1\Private) approximately on line 4 and saved. *Figure 2.23 (b)* can be used as a reference.

```
#include "Misc/FileHelper.h"
#include "HAL/PlatformFileManager.h"
bool UOutputResults::SaveArrayText(FString SaveDirectory, FString
FileName, TArray<FString> SaveText, bool AllowOverWriting = false)
{
SaveDirectory += "\\";
SaveDirectory += FileName;
if(!AllowOverWriting)
{
if(FPlatformFileManager::Get().GetPlatformFile().FileExists(*
SaveDirectory))
{
return false;
}
}
FString FinalString = "";
for(FString& Each :  SaveText)
{
FinalString += Each;
FinalString += LINE_TERMINATOR;
}
return FFileHelper::SaveStringToFile(FinalString, *SaveDirectory);
}
```

Save this, then return to UE4 to build, save, compile.

It is essential to save both tabs and compile in UE4; should any spelling/syntax errors have been made this will appear in the Message Log. This will identify the error and where it is. This coding is shown in *Figure 2.23*. It is good practice to include this coding to ensure the vital information is stored as a variable after each trial completes, then the variable can be saved as a readable file once the experiment has ended. This is explained in more detail in *Section 2.7.1*.

The data can then be stored as a range of file-types, depending on the suffix inserted into the filename. For example, in *Figure 2.23* the data file is saved as comma separated

values (.csv) but a text document (.txt) could also be used. The decision would be based on the programme used to analyse the data after collection has ended; here a .csv document was selected as analysis was to take place in MATLAB. The name of the file is here linked to the `Now` node (see *Section 2.8*)

Instruction: save data before experiment ends

Name the data file appropriately.

```
Level Blueprint > Event Graph > *right-click on empty space* >
*search and select 'Now'* > *right-click 'Now Return Value'* >
Split Struct pin* > *right-click on empty space* > *search and
select 'Append'* > *right-click on 'Append'* > Add Pin *(x4)* >
*type "data-" in 'Append A'* > *type "-" in 'Append C'* > *type
".csv" in 'Append F'* > *link the 'Day', 'Hour', 'Minute' outputs
from Now to 'Append B', 'Append D', 'Append E'
```

The new node should be added into the flow before the game finishes, as in *Figure 2.16*.

```
*right click on empty space on Blueprint* > *search and select
'Save Array Text'* > *link 'Append Return Value' to 'Save Array
Text File Name'* > *drag from 'Save Array Text Save Directory',
release mouse click* > *search and select 'Get Project Directory'*
> My Blueprint > Variables > *drag and drop 'ResultsOutput'* > Get
ResultsOutput > *link 'ResultsOutput output' to 'Save Array Text
Save Text'*
```

The experiment should save before ending, this will be connected to the flow for checking the total number of trials has been completed, which included the `Event Tick` and `Branch`.

```
*link 'Branch True Exec' to 'Save Array Text Exec'* > *right-click
empty space* > *search and select 'Q' from 'Keyboard Events'* >
*link 'Q Released' to 'Save Arrat Text Exec'* > *drag from 'Save
Array Text Exec', release mouse click* > *search and select 'Quit
Game'*
```

2.8 Odds and Ends

Below are some keyboard shortcuts and notes on nodes which may come in useful when creating more complex experiments.

Changing the default scene root: in a blueprint you can get rid of the visible centre point of any object and make it easier to handle by:

`Viewport > Components > Add component > Scene > *click and drag scene on top of Default Scene Root*`

Duplicating assets can be done in three ways once the the object is selected: (1) the standard copy and paste by using 'Ctrl + C' and 'Ctrl + V'; (2) pressing 'Ctrl + W'; (3) holding 'Alt' while using the mouse to click and drag it in a direction.

A keyboard shortcut to rename an asset is to press 'F2' while the asset is selected.

When positioning objects on top of one another, a quick shortcut to do this is to select the object in the environment and position it vaguely above the ground surface; while still holding the left mouse button the 'End' keyboard button can be pressed to automatically shift it down to be level with this surface.

The `Flip-flop` node can be connected to two outputs so that each time the experiment gets to that point, the path will switch between the two outputs. This is useful for having a light turning off and on, or the start point of a stimulus alternating between two positions.

`UtcNow` returns the date and time as it is in Coordinated Universal Time (the same as Greenwich Mean Time), whereas `Now` returns the date and time in your current location. These are used in *Figure 2.16* to name the data file in accordance to when it was created, the 'Return Value' pin has been split to gain access to the measurements individually.

Using a random generator is useful but it is essential to understand the nature of the random item you wish to create. For example; *float*, *integer*, *Bool*, *vector* are all able to

be created using the `Random (appropriate word)` node.

`Set Hidden` should be used if stimuli within the environment are required to become invisible during a portion of the experiment. This can be used for something as simple as a fixation point within the scene, or when virtual buttons are to be shown at the end of trials.

If an experiment is programmed in such a way that different levels are the different conditions, it is possible to seamlessly switch between these using the `Open Level` node. This node can be incorporated into a `Sequence`, where first the conditions of the initial level are set, and then once completed, the experiment can automatically progress to the next level without experimenter interaction.

The `Event Level Reset` sends the execution signal when the level restarts, and a can only be used in the `Level Blueprint`. This is useful when you need something to be triggered once a level has reloaded, an example where this would be useful is if a participant was not positioned in the correct physical area for some trials and so the level does not need to reload but simply restart. This should be connected to a keyboard button press (*Section 2.4.2*) which the experimenter can press if necessary and not have to interrupt the participant.

`Quit Game` should be inserted into the `Level Blueprint`, perhaps after a `Branch` that sets the requirement for ending the experiment after a set number of trials or each condition has been completed. An alternative (and good practise to include for debugging) would be have this connected to a button press, so that the game can be ended easily by the experimenter. An example of this is shown in *Figure 2.16* where the game is programmed to end when either all conditions and trials have been completed, or when 'Q' on the keyboard is pressed.

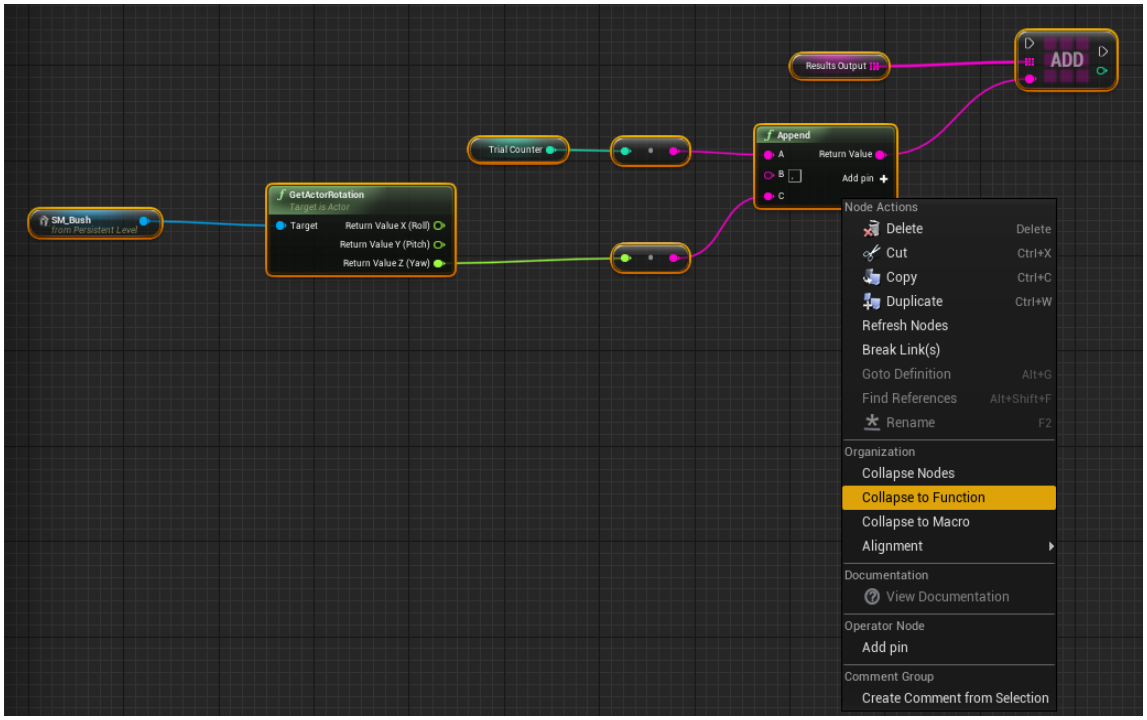


Figure 2.24: An array is made to store the trial number and Z rotation of the `SM_Bush`, these are collapsed into a Function.

Collapsing a section of code is useful for keeping blueprints neat but there are different ways this can be achieved. By selecting multiple nodes, right-clicking on them and then selecting to `Collapse Nodes` allows for a single node representation of the selected section. This has no added features other than making the code tidier and is shown in *Figure 2.18*. Alternatively, after right-clicking, the option to `Collapse to Macro` is very similar, but if this is then copied and pasted elsewhere in the blueprint, any modifications made to one version will appear in all of the duplicates. `Collapse to Function` on the other hand has different properties: this does not allow for latency, so no delays will work within a function. This is shown in *Figure 2.24*. The benefit of using this instead of a macro or collapsed node set is that a function can be called by another blueprint; if for example a timer was created in the `Level Blueprint`, this timer could then be called by the `MotionControllerPawn` blueprint and record reaction time for some controller collision through a visible object.

Instruction: collapse results saving flow

In the `Level Blueprint`, locate the nodes where an array for raw data is created and collapse them as shown in *Figure 2.24*.

```
Level Blueprint > Event Graph > *click on empty space and drag to
select all nodes in this flow* > *right-click on one of the nodes*
        > Collapse to Function > *rename ("RecordData")
```

An additional tip for good practice is to have keyboard inputs as well as controller inputs for events, so that the creator can check events are happening as intended when previewing. As shown in many of the figures (2.13, 2.15, 2.16, 2.21), there are keyboard inputs as well as the controller button presses, so that when piloting it is possible to quickly check the flow operates appropriately.

Previewing what you have built is a useful feature of this program, as it allows you to quickly test and adapt the experiment. By pressing the `Play` button above the `Viewport` you can preview the game in a variety of different ways. The drop-down arrow on this button shows the different ways you can observe the preview; to quickly check you can play within the `Viewport Window`, or for a more accurate representation of what a participant will experience you can play within the VR headset you have connected. Simply press 'Esc' on the keyboard to exit from the preview.

When creating an experiment it is common to encounter errors, or realise events aren't happening the way they were intended. Debugging in UE is simple, right-click on the node which might be causing an issue and select '`Add breakpoint`'. To remove is the same, click on the node again but select `Remove breakpoint`; or if you have multiple points to delete it is quicker to open the `Debug` menu at the top of the window and then `Delete All Breakpoints`. By adding a breakpoint and previewing the experiment, it will pause when the flow reaches this point. In order to view both the blueprint and the experiment at the same time, once any breakpoints have been added you can preview

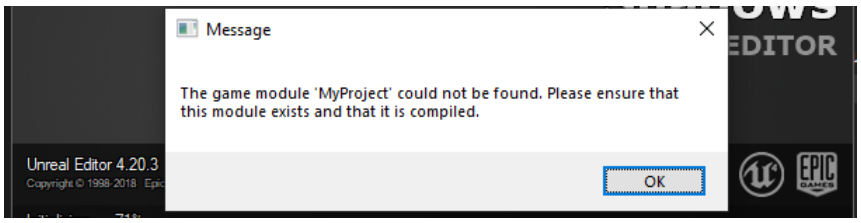


Figure 2.25: Error which may occur when the experiment has been exited before being compiled.

in a `New Editor Window` and move this new window to the side of the blueprint, before clicking on the window and running through the level until the breakpoint is reached.

It might be the case that an error (*Figure 2.25*) appears when trying to open the project. This specific error can occur when certain changes have been made but the project has not been compiled properly. Navigate to the folder this project is saved in through your file explorer, right-click the .uproject file and select 'Generate Visual Studio project files'. This will create a Visual Studio Solution (.sln) in the same folder. Open this and then follow these steps in Visual Studio:

```
Solution explorer > Solution experiment1 *(the project name)* > Games >
                    experiment1 > *right-click* > Rebuild
```

then open with UE4 normally, remember to always save and compile!

When multiple levels are being saved into the project, the level to be loaded as default on startup can be set through:

```
Edit > Project Settings > Project > Maps & Modes > Default Maps >
                    *select 'PsychophysicsTest1'*
```

Adding comments into the blueprints is a way to visually organise different sections of nodes. In *Figure 2.1*, the `Level Blueprint` which can be created using this tutorial is presented, and coloured comments are used to identify which groupings of nodes are controlling similar roles. To achieve this, left-click and hold on an empty are of the blueprint background, then drag to select all nodes and press 'C' on the keyboard, before

releasing the mouse. A pop-up will then appear, where the title and colour can be adjusted. Size adjustments can be made to this comment box by dragging from any corner.

If anything does not work as expected, compiling the project, exiting Unreal Engine and restarting may solve it.

Instruction: loop experimental procedure

Connect three flows created: the `Record Data` function node, the flow where the `Trial Counter` adds 1, and the flow stemming from the 'Z' keyboard press. This is depicted in *Figure 2.11*

```
Level Blueprint > Event Graph > *right-click on empty space in
Blueprint* > *search and select 'N' from 'Keyboard Events'* > *link
'N Released' to 'Record Data Exec' > *right-click on empty space in
Blueprint* > *search 'motion face'* > MotionController (R)
FaceButton1 > *link 'MotionController (R) FaceButton1 Released' to
'Record Data Exec'* > *link 'RecordData Exec' to 'Set Exec' (for
the Trial Counter)* > *return to World Outliner and ensure
'Stimulus' is selected* > *return to Level Blueprint* >
*right-click on empty space on Blueprint* > *search 'visible'* >
Set Actor Hidden In Game > *select 'Set Actor Hidden In Game New
Hidden'* > *link 'Set Exec' to 'Set Actor Hidden In Game'* > *link
'Set Actor Hidden In Game Exec' to 'Delay Exec' (from previous
flow)*
```

2.9 Exporting the Application

When the experiment is complete and ready to be exported, it is important to make sure that all the levels are saved, compiled, and built. These are all options in the `Toolbar` which can be selected before packaging the experiment into a stand-alone application.

Once everything has been finalised, the game can be packaged into a stand-alone application, where it is possible to transport to other PC devices. When developing for untethered, standalone headsets, it is necessary to compile and transfer the application to the headset. As these headsets currently function as Android devices, it is necessary to integrate the development with the Android SDKs. This is beyond the scope of the

present tutorial, which focusses on the creation of experiments for headsets operating via a PC.

Instruction: package project

Check the configurations are set for producing a final version of the experiment and export it.

```
File > Package Project > Supported Platforms...  > *select
                    appropriate platform*

File > Package Project > Build Configuration > Shipping

File > Package Project > *select appropriate platform* > *select
  location for application to be exported to* > Select Folder
```

2.10 Conclusions

UE is a highly versatile programme that gives a level of artistic quality which is much harder to produce in other typical programmes within psychophysics and other research domains, while still accounting for accurate collection of metrics and implementation of parameters.

By following this guide, an experiment that collects quantitative data will have been made, along with some general tips for using UE4 when coding an experiment. By no means is this advice coming from an experienced programmer, just a student trying to share useful information in a concise manner.

Chapter 3

Size and Shape Constancy in Consumer Virtual Reality

The fundamental vision concepts which provide the basis for the following experimental chapter, and the way Unreal Engine has been used to create the experiments presented throughout this thesis have been detailed in *Chapters 1* and *2*. With the increase in popularity of consumer virtual reality headsets, for research and other applications, it is important to understand the accuracy of 3D perception in virtual reality. In this chapter, the perceptual accuracy of near-field virtual distances was investigated using a size and shape constancy task, in two commercially available devices. The HTC Vive and the Oculus Rift were two of the leading virtual reality headsets at the time of publication, and by programming the same experiment to be tested in both devices it was possible to calculate the difference in the general perceived error of distances in these leading consumer devices, and to compare this with result from carefully controlled psychophysical equipment used in other literature.

3.1 Introduction

3.1.1 Distance Perception

Accurate visual perception of our surroundings is important for us to interact successfully and efficiently with our environment. To do this we must be able to estimate the 3D location, shape, and size of objects. Robust measures of perception of these properties in the physical environment have been developed. For example, accuracy in distance perception has been studied using distance bisection (Rieser et al., 1990), verbal estimates (Klein et al., 2009; Mohler et al., 2006), dihedral angle adjustments (Watt et al., 2005a), and blind walking tasks (Knapp, 2003; Kuhl et al., 2006). Accuracy in shape and size perception has been examined by allowing observers to compare a seen object to a fixed standard, such as a circular cylinder (Glennerster et al., 1996; Johnston, 1991; Scarfe & Hibbard, 2006), or to a hand-held object (Brenner & van Damme, 1999). These techniques have been implemented in physical space, using either real or computer-generated stimuli, in order to understand the accuracy of human 3D perception using measures of precision and bias. Precision here refers to the variability across multiple estimates for the same stimulus, while bias refers to any systematic deviation from veridical estimates. These methods have also been adopted in virtual space using 3D computer set-ups, and it has been found that irrespective of the technique used, there does appear to be a mis-perception of distance which may present itself as over-estimation of near distance, under-estimation of far distance, or a combination of both (Chalmers, 1952; Foley, 1991; Higashiyama, 1996; Jenkin, 1957; Johnston, 1991; Kline & Witmer, 1996; Li et al., 2011; Patterson et al., 1992; Scarfe & Hibbard, 2006; Sinai et al., 1998; Viguier et al., 2001; Yang & Purves, 2003).

3.1.2 Applications of Consumer Virtual Reality

As a method of measuring properties of spatial perception, there are significant benefits of using virtual reality (VR), as it allows precise control of stimulus factors that

may be confounded, difficult or impossible to manipulate in real situations. The ability to separate and test these factors individually allows for a better understanding of the visual system's processes during the perception of distance. A striking example of this is the expanding virtual room used by Glennerster et al. (2006), where the scale of the virtual environment increases and decreases as the participant navigates through it. These experiments have found that participants are surprisingly unaware of large changes in scale, and allow for the assessment of the role of binocular and motion cues in 3D perception for a moving observer. These changes in scale of the environment, contingent on the participant's movement, are only possible in VR.

Because of its versatility, VR can be used as a research tool in many fields. In psychology, this includes research into visual perception and spatial cognition (Creem-Regehr et al., 2015; Scarfe & Glennerster, 2015), social interactions (Pan & Hamilton, 2018) and the understanding of conditions such as autism (Kandalaft et al., 2013) and schizophrenia (Freeman, 2008); as well as applications in the fields of education (Freina & Ott, 2015) and therapy (Adamovich et al., 2009). It is now also very affordable, easy to obtain and set-up, and is therefore a viable option for many practical applications, for example in training, design, remote working, and the arts. In all applications, within and beyond research, it is important to establish the accuracy of distance perception within VR if it is to be used successfully. Since errors in distance perception have been found across many tasks and set-ups, including in natural real-world viewing, it is important to know how these errors arise; the extent to which they are exacerbated by properties of VR hardware and software; and the extent to which mis-perception is specific to particular circumstances. For practical reasons, knowing the extent of any over/under-estimation of distance allows for the issue to be mitigated. This is true for both practical applications and for the use of VR in experimental psychology. In studies that address the way in which visual and other cues are used in the perception of 3D space, it is important to understand the nature of any potential biases introduced by the use of VR. It is equally important to know that the virtual environment presented

to the participants is perceived as intended by the experimenters. The purpose of the current study was to quantify the accuracy of shape and size settings in consumer VR, which depend on the perception of absolute distance.

Here, the accuracy of shape and size constancy was measured in two consumer VR systems: the Oculus Rift and the HTC Vive. The factor which defines these as being *consumer* devices is that the target audience is individuals who would use them for leisure, business, education or health-care purposes for example, in contrast to being a dedicated research facility. VR systems for research purposes may have a precise and extensive set-up process, often also accompanied by calibration for the individual observer. Ensuring factors such as the inter-ocular distance, viewing position, lens effects, screen resolution and accommodation distance are all correct can be a time-consuming task. Those using these devices for leisure or non-research applications are unlikely to be able to spend as much time ensuring these features are accurate, and thus consumer devices should be examined for robustness in the absence of these calibrations. Equally, consumer-type VR is increasingly common in all fields of psychological research, and it important to establish that it can be used with confidence without the need for long and laborious calibration procedures.

3.1.3 Biases in Distance Perception

There are many sources of depth information in the physical world and in VR, including both monocular and binocular cues. Monocular cues are those which depend on the information available from one eye. These include simple features such as size, brightness, texture and perspective, as well as motion parallax. Motion parallax can be useful in determining the absolute distance to an object (Bishop, 1989). Binocular cues are those which utilise the information from both eyes, such as binocular vergence and retinal disparity. When an observer is fixating an object, the difference in the direction of the gaze

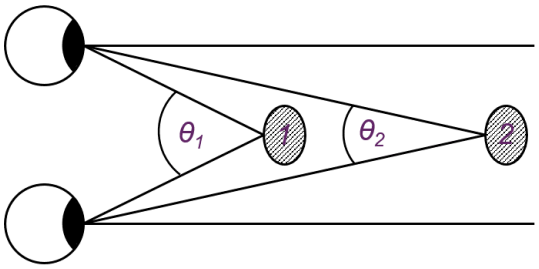


Figure 3.1: The vergence angles of the two eyes when focussing on two different objects: the angle θ_1 is bigger for object *1* than angle θ_2 when converging on object *2* at a further distance.

of the eyes, known as *binocular vergence*, can be used to determine absolute distance. When fixating on a close stimulus, the vergence angle is larger than when looking at a stimulus which is further away (*Figure 3.1*). Binocular disparity refers to the differences in the two retinal images for a given vergence state; these differences relate directly to the 3D structure of the environment, and are therefore a valuable source of depth information.

Viguier et al. (2001) investigated the relationship between the cues of vergence and accommodation, and the perception of distance in near space with a series of tasks, including verbal estimates, bisection, and equal- and double- distance settings in real space. They found that observers under-estimated the distance of targets beyond 50 cm, but for closer targets distance perception was accurate; the authors noted that this range corresponds to that of arm's length, a range within which accurate distance perception is needed for effective motor co-ordination (Paillard, 1991).

In the context of VR, it is important to test which factors might contribute to additional inaccuracies in perception, over and above the biases found in the natural environment. Bodenheimer et al. (2007) compared distance judgements via a bisection task in real space, real space with field-of-view-restricting goggles, and in virtual replicas of real space. The idea here was that a restricted field of view may be one of the factors contributing to distance mis-perception in VR. Two distances (15 and 30 m) were tested,

so that the true bisection results should have been 7.5 m and 15 m, respectively. Each type of environment (real or virtual space) had an outdoor and indoor setting, tested to increase reliability and external validity of the results. Participants were least accurate in the virtual environments. Accuracy was higher for the closer distance than far: if the midpoint was set to be closer to the observer than the actual midpoint, this would indicate that observer over-estimated the near distance, relative to the far space, due to the participant thinking they had set it to be further away than they actually did. The authors describe this as *foreshortening*. This was found in the 30 m environment, but the opposite trend was found in the 15 m environment: the midpoint was set further away than accurate performance which is indicative of a perceptually expanded far space section between the mid-point marker and the end stimuli (referred to as *anti-foreshortening*). In addition, estimates in the natural viewing condition of real space were more precise, and the greatest under-estimation of the mid-point overall occurred in the virtual environments. These results are slightly different from those found in a previous study in physical space (Purdy & Gibson, 1955), however these discrepancies could be due to methodological differences. The authors concluded that a reduced field of view in head-mounted displays (HMDs) was not the cause of the mis-perception of distance. This was supported by Knapp and Loomis (2004) who restricted the vertical field-of-view and found this to have no effect on distance estimation, but instead concluded that this error in distance perception may be the result of a number of relatively subtle factors, including the display resolution, dynamic range, and rendering fidelity, combining to produce an overall composite error.

The results from distance bisection tasks and other experiments (Creem-Regehr et al., 2015) show that in certain areas of space, distances appear to be either under- or over-estimated. If distance is mis-perceived, then a corresponding mis-perception of shape and size is also expected. Consequently, measuring the perceived shape and size of stimuli is a method of indirectly estimating perceived distance. Constancy refers to the perceived properties of an object, such as its size and shape, remaining unchanged

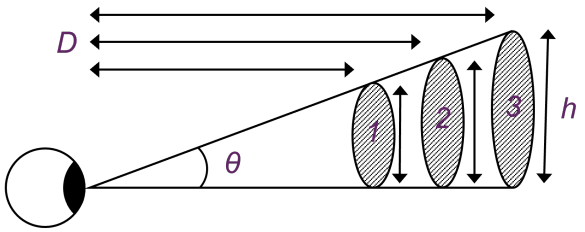


Figure 3.2: Eye observing three stimuli, where θ is the same viewing angle for each; h are the heights of each the object; d are the distances at which each object size would result in the same retinal image. The d to *1* is shorter than the d to objects *2* and *3*, but because the size of *1* is smaller than the others also, it results in the same viewing angle for each object at each distance.

when there are changes in viewing conditions such as its distance from the observer. If a physical ball is moving through space, or is first shown at a close and then a far distance, with unchanging physical dimensions, then the perception of these dimensions should match. However, there is evidence for a mismatch of the perception of these qualities and the true values in both real space and for binocularly-viewed, computer-rendered stimuli. As can be seen in *Figure 3.2*, multiple different combinations of object size and distance will result in the same 2D image being projected onto the retina. If the distance is not known, or is mistaken, then the properties of the object may be mis-perceived, due to the infinite combinations of size and distance which result in the same viewing angle. If the object's dimensions are wrongly identified due to a mis-perception of its distance, it is possible to work out the distance at which the retinal images perceived by the participant would correspond to the correct physical object dimensions. This approach can be referred to as a *constancy task* and is an indirect method of assessing distance perception. One method of testing this phenomenon involves setting a dimension of a stimulus to match that of another object, which can be compared visually or haptically, across a number of distances. For example, an observer may be asked to set the apparent shape and size of a viewed ellipsoid to match a hand-held tennis ball (Brenner & van Damme, 1999). If true constancy was achieved then the same radius settings would be made across distances, as shown in *Figure 3.3* (black line in *Radius Predictions*, slope of zero). Translating this into the effective distance observed would

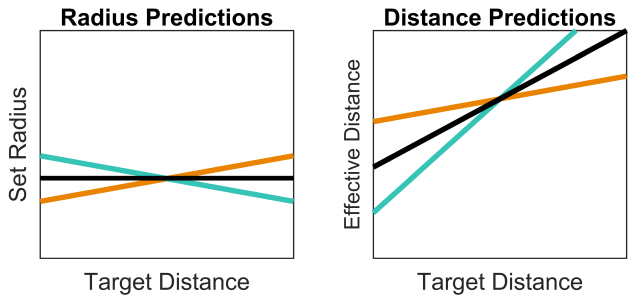


Figure 3.3: Two sets of possible trends in how setting might vary with distance, and the interpretation of these in terms of mis-perception of distance. The *Radius Predictions* shows the trend of true constancy in black, with an intercept of what the radius of the physical stimuli is, and slope of zero; under-constancy, the prediction based on prior research, is shown in orange, in which settings are made too small at close distances and too large at far distances. The opposite possible trend, of over-constancy, is shown in turquoise, with setting being too large at close distances and too small at far distances. The *Distance Predictions* shows the possible trends for the effective distances from the radius settings with accurate perception of distance having an intercept of zero and slope of one in black; under-estimation shown with a slope of less than 1 in orange and over-estimation with a slope of greater than 1 in turquoise. Predictions have been made with a non-zero intercept, since accurate settings tend to be made at an intermediate distance.

cause the perceived and actual distances to be the same (black line in *Distance Predictions*, slope of one). A different possible outcome would be a progressive over-setting of the radius, such that it increases with presentation distance (shown by the orange line). This would mean that the distance perceived increases at a lesser rate with the distance presented than it should. The opposite possible result is shown by the turquoise lines; progressively smaller settings of the radius with distance would be consistent with perceived distance increasing with the presented distance at a greater rate than it should. Evidence from previous experiments (Brenner & van Damme, 1999; Glennerster et al., 1996; Johnston, 1991; Scarfe & Hibbard, 2006) has typically found under-constancy: perceived distance does increase with presentation distance, but at a lesser rate than expected under veridical perception, as indicated by the orange line in *Figure 3.3*. In this figure, both under- and over-constancy outcomes have been plotted with a non-zero intercept. This is because participants tend to perceive intermediate distances accurately.

Using a physical reference object, which cannot change in its properties (size or shape), has the advantage of providing an external reference, experienced through a different sensory modality, which is therefore not affected by bias in the visual perception of 3D space. This technique has been shown to be a robust measure of distance mis-perception (Brenner & van Damme, 1999). In a carefully calibrated, 3D screen display, participants were presented with ellipsoidal surfaces. Participants adjusted the size and depth of the ellipsoid, until it matched a tennis ball they held in one hand. They then held the tennis ball out in front of them, to indicate its perceived distance. As presentation distance increased, so did indicated distance, although indicated distance was gradually more under-estimated with increasing presented distance. As the distance of the virtual ellipsoid increased, the settings of size and depth also increased, which the authors concluded was due to participants mis-estimating the distance to the stimuli. These findings support the predictions shown by the orange lines in *Figure 3.3*. The scaling technique showed an over-estimation of perceived space at the near-distance trials, and an under-estimation of perceived space at the further distances. This is consistent with the previously discussed compression found to occur in a range of different environments, including the real world (Smallman et al., 2002) and when using HMDs (Plumert et al., 2005; Ponto et al., 2013).

3.1.4 Shape and Size Constancy

The geometry relating perceived size and distance to the retinal image is shown in *Figure 3.2*. The retinal size of the image, θ, is the same for each object, shown here at three distances. The relationship between θ and the distance to the object ($\mathtt{D}$) and its height ($\mathtt{h}$) is given by:

$$\mathtt{tan}(\theta) = \frac{\mathtt{h}}{\mathtt{D}} \tag{3.1}$$

In this experiment, the same task as Brenner and van Damme (1999) was used, in

which observers were required to match the shape and size of a virtual ellipsoid to those of a hand-held tennis ball. It is possible to distinguish between the *rendered* size and distance of the object ($\mathtt{h}_{\mathtt{rendered}}$ and $\mathtt{D}_{\mathtt{rendered}}$), as geometrically simulated in VR, and the corresponding *perceived* properties ($\mathtt{h}_{\mathtt{perceived}}$ and $\mathtt{D}_{\mathtt{perceived}}$). These are related to the retinal image size, and to each other, as follows:

$$\mathtt{tan}(\theta) = \frac{\mathtt{h}_{\mathtt{rendered}}}{\mathtt{D}_{\mathtt{rendered}}} = \frac{\mathtt{h}_{\mathtt{perceived}}}{\mathtt{D}_{\mathtt{perceived}}} \tag{3.2}$$

This can then be rearranged to express the perceived distance in terms of the rendered distance and size, and perceived size:

$$\mathtt{D}_{\mathtt{perceived}} = \frac{\mathtt{D}_{\mathtt{rendered}}\mathtt{h}_{\mathtt{perceived}}}{\mathtt{h}_{\mathtt{rendered}}} \tag{3.3}$$

This transformation allows for the perceived distance to be calculated, assuming this to be the source of the error in perceived size. This is achieved by using the known values of $\mathtt{D}_{\mathtt{rendered}}$ and $\mathtt{h}_{\mathtt{rendered}}$ and assuming that $\mathtt{h}_{\mathtt{perceived}}$ accurately matched the felt size of the hand-held reference ball.

The left and right eyes' images when viewing an object depend on the observers' inter-pupillary distance (IPD) as well as the 3D shape and location of the object. In the presented experiments, the IPD of individual participants was not measured, and a single value of 63 mm, taken to represent the mean of the population (Dodgson, 2004), was used for all participants. This decision was made to reflect that in most research and other applications, the IPD was not measured or controlled for (see *Chapter 6*), and the purpose of this study was to measure the accuracy of 3D vision under typical conditions. Even in situations in which the viewing geometry is carefully measured and calibrated, the IPD may be fixed at an average value, so that it conflicts with that of most observers (Glennerster et al., 2006).

In both headsets there is the option to minimally adjust the IPD (stepped and not continuous control), however this was kept constant throughout the trials because

casual users are unlikely to adjust these correctly to their own measurements (few if any members of the public will accurately know their own IPD) and some consumer HMDs on the market do not allow for this manipulation. If anything, users might adjust the headset IPD to maximise comfort, but this does not guarantee that the set IPD will match their physical IPD. For example, observers may simply set just enough binocular parallax to create a stereoscopic effect, so called *micro-stereopsis* (Siegel & Nagata, 2000). By keeping the IPD constant, the results obtained would be generalisable to other headsets with similar specifications and to the users who do not know and/or correctly adjust to their own measurements. The effects of a mismatch between the actual and assumed IPD can be calculated geometrically, and depends on the effective distance of the screen from the participant's eyes (Drascic & Milgram, 1996; Hibbard et al., 2020). In practice, however, the effect on perception is expected to be much reduced in comparison with geometric predictions, since vergence is a rather poor cue to absolute distance (Collewijn et al., 1991). For example, changes in perceived distance with vergence distance show a very shallow slope (Foley, 1980; Johnston, 1991), and a great deal of variability between participants (Gogel, 1977; Morrison & Whiteside, 1984; Richards & Miller, 1969). It has been proposed that, rather than using vergence as an absolute cue to distance, changes in vergence are used as a cue to relative distance, or that the use of vergence as an absolute distance cue is poorly calibrated (Brenner & Van Damme, 1998). As such, biases in distance arising from the mismatch between assumed and actual IPD are likely to be much less severe than predicted geometrically, leading to greater consistency between participants than might otherwise be found. Finally, even with complex photometric calibrations methods used in a lab setting, the calibration is carried out for a single camera separation; for example a single IPD (Gilson et al., 2011; Scarfe & Glennerster, 2019). Therefore, using a fixed IPD in the present experiment allowed better generalisation to lab-based studies and everyday use.

Similarly, it is possible in principle that biases in the geometrical rendering of the stimuli could have been introduced by the hardware or software used, such that the images presented did not accurately reflect the intended projected images. Again, while

this is something that can be measured directly (Gilson et al., 2011), the intention here was to measure the accuracy of perception when VR is used by consumers, or in typical research applications, without this lengthy calibration.

3.1.5 Current Study

The aim of the current study was to measure the accuracy of distance perception in two consumer VR devices using shape and size constancy tasks. Two devices, the Oculus Rift and the HTC Vive, were used in order to assess the generalisability of the collected results. A difference in results obtained from the two devices was not expected to be found, more so this was a replication in two of the main competing consumer VR devices available on the market. Whilst, on the face of it, the specifications of the headsets are similar, there are in fact some important key differences which could affect the perception of distance, depth, and shape. For example, the lens systems of the two headsets differ. Additionally and anecdotally, industrial users who have rendered large warehouse-size rooms in VR have reported users experiencing differences in the perception of scale[1]. Thus, whilst a difference between headsets seemed unlikely, it was a possibility. Comparing two headsets resulted in collecting twice the amount of data, but it allowed for (1) comparison two of the main competing consumer VR devices available on the market (rather than less widely adopted devices); and (2) replication of findings across the two headsets.

An ellipsoid was shown to participants at randomised distances within near-space; the task was to match the shape and size of the visual ellipsoid to that of a tennis ball, using buttons on a controller to change the depth and size. Based on previous research, it was expected that as presentation distance increased, so too would the size and depth settings, consistent with a progressive under-estimation of distance. Furthermore, it was expected that size and depth measurements would be positively correlated, if variation in these results from the same underlying mis-estimation of distance. Mis-estimates of shape and size were used to infer the accuracy of distance perception in the two consumer

[1]Peter Scarfe, personal communication

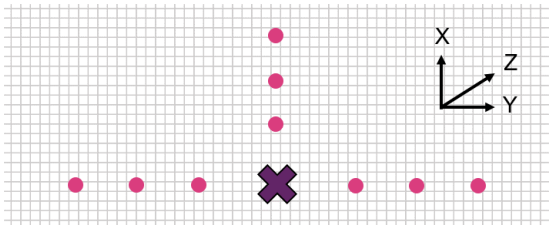


Figure 3.4: Bird's-eye view of the correct position for a participant to position themselves (marked as the cross) in the virtual environment, triangulated against visible references. The closest of each of the three sets of balls would occlude the further two in each set, when the participant is positioned in the correct location.

VR systems. The overall goal was to assess the degree of shape and size constancy in consumer VR in comparison with the results of previous psychophysical experiments. Thus, while a degree of under-constancy is expected in both shape and size perception, the aim here is to assess the extent to which this might be exacerbated by the use of consumer VR, and the implications this might have for research that makes use of VR.

3.2 Methods

3.2.1 Participants

Opportunity sampling of students from the University of Essex was used to recruit 40 participants, all of whom were naive to the background and hypotheses of the study. The age ranged from 18 to 49, with a mean (standard deviation) of 21.8 (5.6), with 25 males and 15 females. The average completion time was 45 minutes and all participants had normal or corrected-to-normal vision.

3.2.2 Apparatus

The experiment was conducted in two VR devices - the HTC Vive and the Oculus Rift - and ran on a PC with an Nvidia GeForce GTX 1060 graphics card. The IPD was set to the average of 63 mm (Dodgson, 2004). The associated Rift Touch controllers and Vive Motion controllers were used by participants, along with a standard tennis ball with a 3.5 cm radius. Display specifications of the two headsets are summarised in *Table 3.1*.

Table 3.1: Display properties of the two VR headsets used in this chapter.

	HTC Vive	**Oculus Rift**
Field of view (H x V)	110 x 113°	94 x 93°
Display resolution per eye	1200 x 1800 pixels	960 x 1080 pixels
Pixel size	6.2 arc min	5.2 arc min
Lens	Fresnel	Hybrid Fresnel
Refresh rate	90 Hz	90 Hz

3.2.3 Stimuli

Stimuli were created and presented using Unreal Engine 4.18 and within the application, the X axis represents the horizontal left/right, Z vertical up/down and Y forward/backward.

Either the left or right controller could be used by participants to modify the stimuli, the choice was theirs to make. In the Rift controller the face buttons, A/B on the right controller and X/Y on the left, changed both the width and height (X and Z axis) and the grip and trigger changed the shape of the ellipsoid by scaling the Y dimension of the ellipsoid. On the Vive controller the four D-pad buttons achieved the same results.

One ellipsoid was visible throughout the experiment. A black and white, random dot image was used as the ellipsoid material to provide a surface texture. In each trial the ellipsoid was presented at a randomised distance between 40 to 100 cm from the participant along the Y axis, and the size and depth of the ellipsoid were also randomised to values of between 0.5 and 2 times the starting X, Y, and Z values of 7 cm. The minimum and maximum values that could be set for each dimension were 1 cm and 50 cm, respectively.

All stimuli throughout the experiment were presented at eye-height for the individual participant, on a background with no environmental cues that could have influenced their estimates (Lappin et al., 2006): a clouded sky was rendered in full 360° and the ground plane was untextured. The far distance of the sky means that it was rendered with an

optical angle disparity of zero, while the untextured ground plane provided no disparity information.

3.2.4 Task and Procedure

To ensure that the participant was in the correct position in the room, a brief calibration process was completed before the task: two sets of spheres (with the same visual attributes as the target ellipsoid) positioned along the X axis and one along the Y axis were all aligned so that when in the correct position, the closest sphere in each set occluded the two spheres behind. The participants were required to find the correct area within the room for this, which is shown in *Figure* 3.4. All visible spheres could be moved upwards and downwards until they were at eye-height for the participant. These spheres were then removed by a key-command from the experimenter, leaving only the target ellipsoid visible.

The participant grasped the controller in one hand and the tennis ball in the other, neither were visible to them while wearing the headset. This task was to seen the seen object in the virtual environment to the haptically felt physical reference object. The ellipsoid visible in the headset always had the same Z coordinate as the participant's eye-height which was established during the calibration task. In each trial the participant altered the appearance of the virtual ball to match that of a tennis ball. A press of any of the buttons on the controller would manipulate the size or depth to increase or decrease at a rate of 1 cm per second.

Once the participant decided that the virtual ellipsoid matched the physical ball, a verbal indication to the experimenter was given so that the size, depth and distance measurements could be recorded. Following this, the next trial began and the distance, initial size, and depth were all randomised. This was repeated until 100 trials were completed.

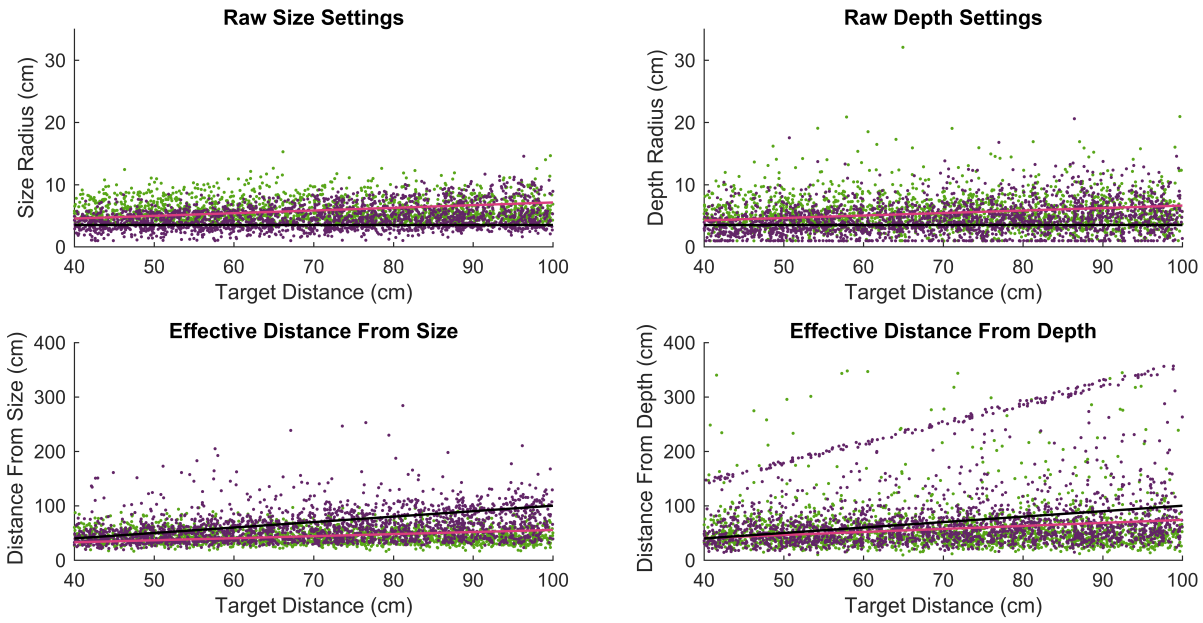


Figure 3.5: Raw settings for all participants, for the HTV Vive (green) and Oculus Rift (purple). Top row shows the actual radius settings for size and depth. The bottom row shows the calculated scaling distance in each case. Pink lines superimposed are the predictions from the linear regression model that was fit to the data, black lines are accurate performance.

3.3 Results

3.3.1 Analysis Strategy

The effective distances at which participants were perceiving the stimuli in each trial were calculated using *Equation 3.3*. With these data, linear mixed effects models were used to analyse the depth settings and the size settings independently. The slope values from these analyses indicate the amount of distances being perceived by participants, for the presentation distances of the target stimulus; the intercept values identify the systematic bias of perception.

3.3.2 Measurement Settings

The raw data, for all participants, are plotted in *Figure 3.5*. The data were initially analysed using one linear mixed effects model for each dependent variable, to assess how size and depth settings were affected by distance. Under perfect scaling, no effect

of distance on these settings would be expected. Based on previous studies, however, imperfect scaling is expected, such that both settings will tend to increase with distance, as both size and depth are progressively under-estimated. Data from the two headsets were combined and analysed with two models, with either size or depth as the dependent variable, and object distance as a linear covariate, headset as a categorical factor, and their interaction, included as fixed effects predictors. Random effects across participants were also included. For each dependent variable, two possible random-effects structures were considered (random intercepts only, or random intercept plus the repeated measures factor of distance). The preferred model was chosen as the one having the lowest AIC. AIC values for both models are shown in *Table 3.2* for each dependent variable. In each case, the model with distance as a random factor produced a better fit than the model with only random intercepts, and this structure was therefore used in both cases. However the drawn conclusions here (the direction, size and significance of the effects) were not affected by this choice of model.

The models are summarised in *Table 3.3* for size settings and *Table 3.4* for depth settings. In both cases, the intercept was less than the expected value of 3.5 cm, indicating a small over-estimation of the radius at near viewing distances. Settings increased significantly with increasing distance. These results represent a failure of perfect constancy, but are consistent with previous results (e.g. Brenner and van Damme (1999), Johnston (1991), Scarfe and Hibbard (2006)). There was no effect of headset, and no headset-by-distance interaction. There was thus no evidence of a difference in distance scaling between the two headsets.

A *post hoc* note was made of participants who moved or tilted their head to the side (while maintaining the original position of the chair) as these could be classified as using motion parallax. It was found that three participants using the Rift and six using the Vive used this additional cue and after running the same regressions as before to compared those who did and did not use the cue, no significant results were found. Using this cue did not aid accuracy in either device. It should be noted that, in this experiment, the role of motion parallax was not directly assessed. However, this can be

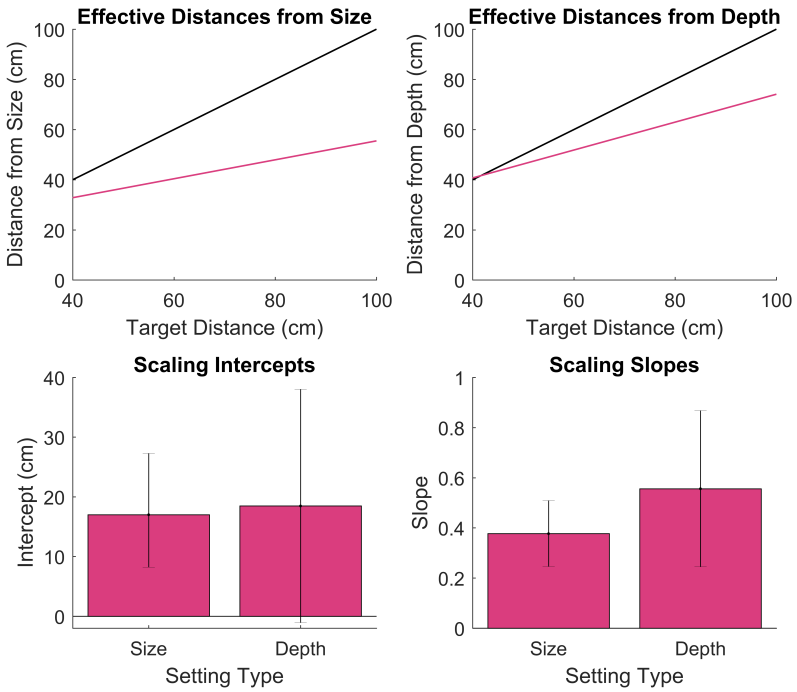


Figure 3.6: Results of the effective distances, from the raw data. Pink lines showing the fitted regression, black lines showing true constancy with an intercept of zero and slope of one. The bar graphs show the mean 95 % confidence limits for each fitted parameter.

achieved both through tracking head movement data from the headset, and experimentally manipulating the degree of motion required of the participants.

3.3.3 Effective Distance

For each trial, the effective size and depth distances were calculated, following *Equation 3.3*. A linear scaling of distance was used for each since, although binocular disparity scales with the square of distance, other cues (such as texture) scale linearly. These were analysed with mixed effects models, again with distance, headset and their interaction as fixed factors, and distance and intercept as random effects. This reanalysis allows us to calculate the degree of constancy in each case. This is presented in *Figure 3.6*, where a zero slope for distance would indicate that participants were not taking distance into account at all in making their settings, while a slope of 1 would indicate perfect scaling. Slopes of 0.38 for size (shown in *Table 3.5*) and 0.56 for depth (*Table 3.6*) were found, indicating incomplete scaling. The distance scaling calculations were performed on the

basis of a linear scaling of depth with distance. For binocular disparity, in contrast with other cues, it is known that perceived depth scales with the square of distance. With this in mind, scaling distances were recalculated to take this into account and the result was a slope of 0.71. While this represents better constancy, it is less consistent with the slope estimated for the perception of size. This is likely to reflect that other cues, including texture, were also present in the stimuli, and these scale with distance, rather than its square. The intercept (the effective perceived distance if the rendered distance was zero) was close to 18 cm in both cases. This combination of intercept and slope mean that close distances tend to be over-estimated, and far distances under-estimated. The crossover distance at which rendered and effective distance were equal was around 30 cm for size settings and 40 cm for depth settings.

3.3.4 Correlation Between Size and Depth

A set of correlations was undertaken on the raw depth and size data from both headsets combined, and a correlation coefficient calculated for each participant. A one-sampled t-test against zero was performed on all correlations for size, and another for depth correlations. The analysis on the size correlations was T (39) = 10.42, p <0.001; for depth T (39) = 9.10, p <0.001. These results show that the variability in the two settings were positively related, as expected if errors in both represent a common mis-perception of distance.

3.4 Discussion

The purpose of the current study was to use shape and size constancy tasks to infer the accuracy of distance perception in virtual environments for two consumer ready, uncalibrated HMDs. Participants wore either the Oculus Rift or HTC Vive and observed an ellipsoid within near-space. Their task was to match the size and depth of the ellipsoid to a physical tennis ball they were able to grasp. Accurate results would show constant settings made across the range of distances presented.

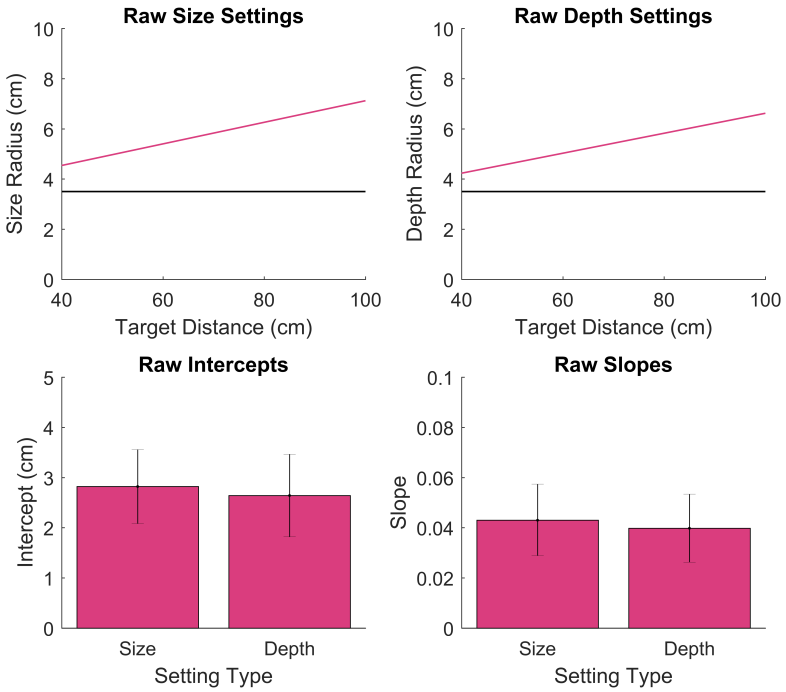


Figure 3.7: Results of the raw settings made by all participants. Pink lines showing the fitted regression lines, black lines show true constancy with an intercept of 3.5 and slope of zero. The bar graphs show the mean 95 % confidence limits for each fitted parameter.

3.4.1 Summary of Findings

Perfect constancy across distances for size or depth was not found in either device. In *Figure 3.7* it can be seen that as the presentation distance increased, both the size and depth settings also increased and these trends did not differ between the devices. Consequently, scaled distances for both size and depth show a general mis-perception of distance in all cases. The results show an over-estimation of near distance and under-estimation of far distance. Overall, the results show a general under-constancy, with changes in effective distance appearing to be 47 % smaller than expected from the variation in rendered distance.

The radius settings of the original data set (depth and size) should be positively correlated, in that as participants set the ellipsoid to be increasing in size (as distance

increased, which was expected due to presumed lack of constancy) they would also increase the depth setting (if the same error affected both settings). Correlations showed that generally this was the case, and a t-test on these showed that the perception of size and depth are related, where the systematic errors made for one are likely to be related to the same underlying cause as the errors in the other.

3.4.2 Comparing to Prior Research

It is important to note that the degree of shape and size constancy, of around 47 %, is very similar to that found for carefully calibrated 3D displays (Brenner & van Damme, 1999) when a similar task was used. Therefore, no evidence for additional sources of mis-perception associated specifically with consumer VR were found. In other studies, distance compression has been found in VR when compared with the real world (Creem-Regehr et al., 2015). These differences may include cognitive effects such as expectations of the room size, rather than any biases introduced by rendering of binocular or motion cues, image quality or field of view (Creem-Regehr et al., 2015). In the current experiment, with very limited visual cues available, compression of the perceived distance range is consistent with the uncertainty of the perceptual information provided (Mon-Williams et al., 2000).

3.4.3 Future Research and Conclusion

This study also provides a starting point for future experiments to examine the role of specific factors in accurate perception in VR, such as the development of high-quality graphics, to improve the 3D experience, and how much these additional features might affect distance perception in immersive HMDs. It has been suggested that the visual system may not always need to recover full metric shape information in everyday interactions (Bradshaw et al., 2000). Investigations into which specific cues, such as perspective, texture, binocular disparity and motion parallax are necessary for producing accurate or aesthetically pleasing displays within VR are therefore important areas for research

using consumer VR. There is evidence that perspective and binocular disparities make different contributions to depth and shape perception (Van Ee et al., 2002; Welchman et al., 2005) and that the cues are differentially used at different viewing distances (Hillis et al., 2004; Keefe et al., 2011; Surdick et al., 1997) due to the way in which their relative reliability changes with distance (Watt et al., 2007). The results and methodological setup here can therefore be used to provide a basis to explore these specific cues and their contribution to distance perception in VR.

Overall, the results here are consistent with the findings of other studies into constancy within 3D set-ups. There were no specific problems identified for distance scaling in consumer VR, with the expected result being found that observers reported objects to be smaller at further distances. Importantly, no systematic differences between the two headsets used were found, and no evidence for errors in constancy over and above what would be expected when sparse cues to distance are available. In conclusion, consumer VR headsets provide a sufficiently high degree of accuracy in distance perception which, despite their low resolution and optical distortion, is comparable to more specialist 3D setups, to allow them to be used more confidently in future experimental vision science, and other applications in psychology.

Table 3.2: AIC goodness of fit measures for the two models (with or without random distance slopes) for the two dependent variables). In both cases the model with random slopes gave a better fit, as indicated by the lower AIC values.

Formula	**AIC**
size ~ distance + headset + distance*headset + (1 | participant)	11808
size ~ distance + headset + distance*headset + (1 + distance | participant)	10591
depth ~ distance + headset + distance*headset + (1 | participant)	17178
depth ~ distance + headset + distance*headset + (1 + distance | participant)	16994

Table 3.3: Results for the size settings. Size settings increased with increasing distance, indicating incomplete constancy. There was no effect of headset, and no headset-by-distance interaction. These results indicate no difference in size constancy in the two headsets.

Predictor	**Estimate**	**SE**	**t**	**DF**	**p**	**Lower**	**Upper**
Intercept	2.821	0.376	7.497	3996	<0.001	2.083	3.558
Headset	-0.359	0.532	-0.675	3996	0.500	-1.403	0.685
Distance	0.043	0.007	5.946	3996	<0.001	0.029	0.057
Headset-by-distance	-0.012	0.010	-1.215	3996	0.224	-0.033	0.008

Table 3.4: Results for the depth settings. depth settings increased with increasing distance, indicating incomplete constancy. There was no effect of headset, and no headset-by-distance interaction. These results indicate no difference in depth constancy in the two headsets.

Predictor	**Estimate**	**SE**	**t**	**DF**	**p**	**Lower**	**Upper**
Intercept	2.642	0.420	6.293	3996	<0.001	1.819	3.465
Headset	-0.600	0.595	-1.008	3996	0.313	-1.768	0.567
Distance	0.040	0.007	5.757	3996	<0.001	0.026	0.053
Headset-by-distance	-0.005	0.010	-0.553	3996	0.580	-0.025	0.014

Table 3.5: Results for the scaling distances calculated from the size settings. Effective distance increased with rendered distance at a rate of 38 %, indicating incomplete scaling. There was no effect of headset, and no headset-by-distance interaction. These results indicate no difference in size constancy in the two headsets.

Predictor	**Estimate**	**SE**	**t**	**DF**	**p**	**Lower**	**Upper**
Intercept	17.756	4.861	3.652	3996	<0.001	8.225	27.287
Headset	7.853	6.881	1.141	3996	0.2538	-5.637	21.343
Distance	0.377	0.067	5.616	3996	<0.001	0.246	0.509
Headset-by-distance	0.13453	0.095	1.415	3996	0.157	-0.052	0.321

Table 3.6: Results for the scaling distances calculated from the depth settings. Effective distance increased with rendered distance at a rate of 56 %, indicating incomplete scaling. There was no effect of headset, and no headset-by-distance interaction. These results indicate no difference in depth constancy in the two headsets.

Predictor	**Estimate**	**SE**	**t**	**DF**	**p**	**Lower**	**Upper**
Intercept	18.474	9.971	1.853	3996	0.064	-1.074	38.022
Headset	9.132	14.139	0.646	3996	0.5184	-18.588	36.853
Distance	0.556	0.159	3.501	3996	<0.001	0.245	0.868
Headset-by-distance	0.188	0.225	0.836	3996	0.403	-0.253	0.629

Chapter 4

Contribution of Differing Visual Cues on Distance Perception in Virtual Reality

An under-estimation of distance has been identified in virtual reality through the previous chapter. As an extension of this, here, the contribution of binocular disparity and the pictorial cues of linear perspective, texture, and scene clutter to the perception of distance in consumer virtual reality were assessed. As additional cues are made available, distance perception is predicted to improve, as measured by a reduction in systematic bias, and an increase in precision. Here, it was assessed (1) whether space is non-linearly distorted; (2) the degree of size constancy across changes in distance; and (3) the weighting of pictorial versus binocular cues in virtual reality. By determining whether specific visual cues significantly enhance performance in virtual environments, the applications of this can be used when developing virtual environments while aiming to minimise the host PC's memory and processing power required to run.

4.1 Introduction

Identifying the distances between oneself and other objects is essential for our everyday actions; from reaching to pick up an object, to perceiving how much leeway there is before stubbing your toe on a table, or cautiously keeping a safe distance from a cliff edge. A broad array of visual information from our environment allows us to interpret where we are, what surrounds us, and what we can do in that space. Cues inevitably vary in both the nature and reliability of information that they provide, and our visual system needs to take this into account in best weighing up the evidence when judging distance. The aim of this three-part study was to determine how visual cues are combined in the perception of distance in complex, naturalistic settings. Virtual reality (VR) was used to precisely control the availability of distance cues, and assessed the perception of distance in terms of both its bias and precision, using distance sectioning and size constancy tasks.

4.1.1 Factors Affecting Distance Perception

Distance estimation is influenced by environmental context, the availability of depth cues, and the task for which it is used (Proffitt & Caudek, 2003; Wickens, 1990). There are many visual cues to depth, and they can be broadly categorised into those that are available via a single monocular image (pictorial cues); those that depend on the differences in the vantage points of our two eyes (binocular cues); those that depend on the motion of objects or the observer (shape from motion and motion parallax) and finally physiological aspects which are guided by distance (such as the accommodation of our lenses, or the binocular convergence of our eyes). In VR, perceived depth may also be affected by *system-related* factors such as the field of view of the display or the accommodation distance required to bring this into focus.

The way different cues contribute to distance estimates has been explored in a variety of different ways. For example, Sinai et al. (1998) explored how texture impacted distance perception in a psychophysical experiment which utilised visually-guided action

tasks and Wu et al. (2004) conducted similar tasks in a series of visually and non-visually guided experiments. Ooi et al. (2001) also used a visually-guided action task but here the impact of angular declination below the horizon was utilised during distance estimations. Throughout these two experiments, participants performed similar tasks but with a different paradigm. From these studies, and the results of many others described previously, it is possible to distinguish the underlying mechanisms of different visual cues. The benefit of having a multitude of objects in an environment is a cue in itself; more specifically, texture gradient and vergence are cues to distance, and binocular disparity for the depth. These examples are explored in the current experiment.

Crompton and Brown (2006) tested egocentric distance estimates in physical space, in industrial and rural settings. Participants estimated the distance to an identifiable structure, and drew increments of this distance on an un-scaled map. It was found that in a rural village, distance was overestimated by a factor of three, compared to an overestimation factor of 1.5 in the industrialised town. Thus, some environmental factors in the scenery must come into play when interpreting scene distances. The study showed that large biases in perceived distance can occur in the real world; testing such phenomena in VR allows us to keep extraneous cues constant while testing the effect of specific cues in separate, yet consistent, environments. This is the method of the current study.

In real, physical environments, it is difficult to separate and control distance cues while still presenting participants with a rich, naturalistic stimulus. This desired combination of complexity and control is easy to achieve in virtual environments, such as stereoscopic displays and VR (Barfield & Rosenberg, 1995; Hibbard et al., 2017b; Hibbard et al., 2017c; Lampton et al., 1995; Surdick et al., 1997; Witmer & Kline, 1998). In VR, it is possible to manipulate the presence of cues and thus assess their importance for the perception of distance. In turn, these findings are important for optimising the display of 3D information in VR. Combinations of cues which are known to improve the accuracy of perceived depth can be prioritised in the rendering of scenes, while those

which have a minimal contribution can be omitted. VR headsets place high demands on graphical rendering, and it is therefore beneficial for software designers and researchers to establish which visual cues can provide for the best perception of space within these virtual environments, so that power is not unnecessarily wasted on additional components that do not enhance the experience.

Kline and Witmer (1996) suggested that pictorial distance cues such as linear perspective, relative size, brightness, and height in the visual plane have lower importance than physiological and system-related (hardware) cues in near space. An experiment investigating these two types of visual cues was conducted, manipulating specifically the field of view (FoV; system-related) and texture type (pictorial). Participants made distance estimates against a wall in a stereoscopic display, with their head position locked. It was found that FoV had more impact than texture at distances up to 12 feet (3.7 metres), and that simply having texture available improved estimates of distance compared to when no texture was present. There was a consistent overestimation of distance in trials with a narrow FoV, and consistent underestimation in trials with a wide FoV. The authors concluded that FoV was most heavily relied upon at near distances because the difference in accuracy was greater between the narrow and wide FoVs, than the difference in texture conditions. For instance, the overestimation of distances in the narrow FoV condition was reduced by the presence of texture, but not entirely.

4.1.2 Methodological Set-up

Psychophysical methods have been used to assess the contributions of binocular and pictorial cues on the perception of distance, using stereoscopic displays. For instance, Surdick et al. (1997) used two-alternative forced choice methods to assess the contributions of multiple cues to the perception of depth. They found that perspective cues were significantly more efficient at conveying depth at a distance of 1 metre than cues such as relative brightness or binocular disparity. In the current study, a similar approach

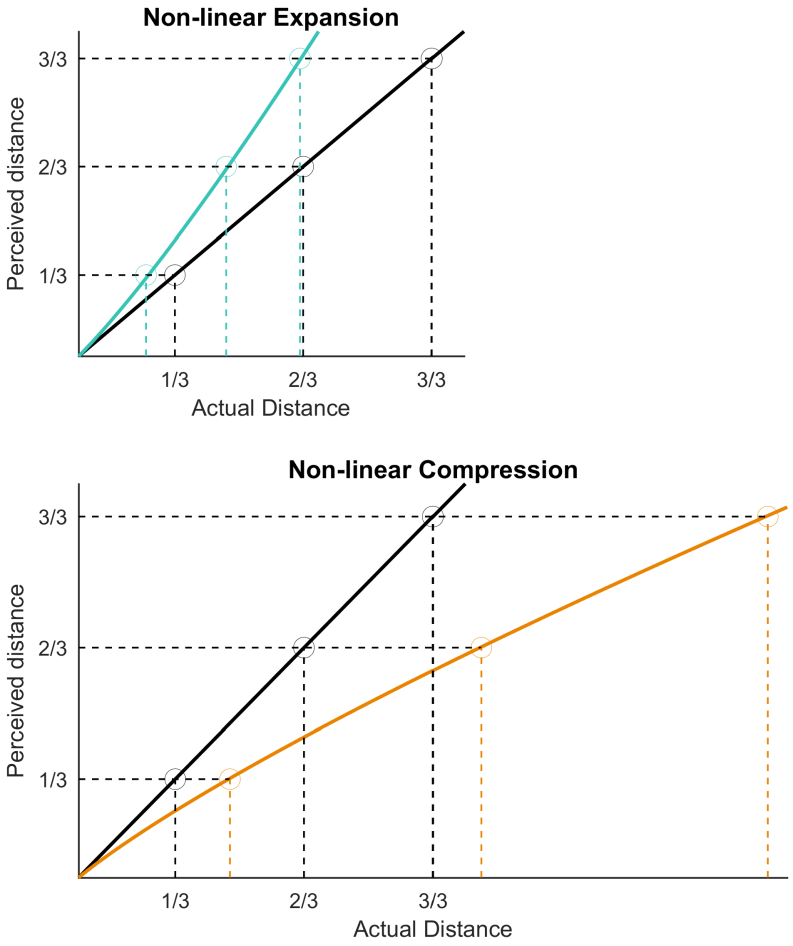


Figure 4.1: Relationships between actual distance and perceived thirds in the Trisection task, presented for two possibilities of non-linear mis-perception of distance. Accurate linear perception in black; perceived thirds to be set not far enough in the expansion instance; perceived thirds shown to be set further than the true thirds in the expansion instance.

is applied in VR, to determine the contributions of binocular and pictorial cues to the accuracy of distance perception under free-viewing of complex, naturalistic displays.

The accuracy of distance perception is characterised by the bias and precision of participants' judgements. Here, bias refers to any systematic errors in judgements, whereas precision refers to the variance of these over repeated judgements. Two tasks were used: distance trisection and size constancy, to assess how these measures vary across distance and cue conditions.

Bisection is a commonly used practise for measuring biases in the perception of

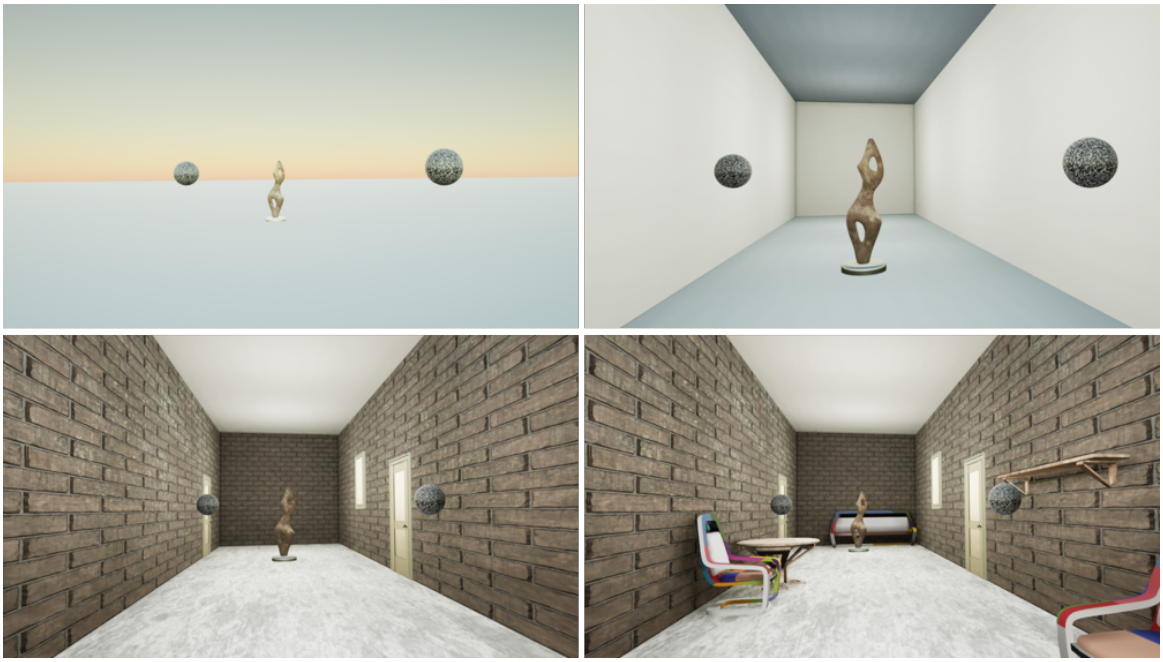


Figure 4.2: The four Trisection task environments: `Sparse` (top left), `Perspective` (top right), `Textured` (bottom left), `Cluttered` (bottom right)

distance (Bodenheimer et al., 2007; Hornsey et al., 2017; Rieser et al., 1990), in which an observer splits a specified distance into equal halves with a marker. This method allows us to determine whether perceived distance is linearly related to physical distance, or whether there is systematic expansion of compression of far space relative to near space. For example, if far space is perceptually compressed relative to near space, then it would be expected that the marker would be set closer than the true midpoint, to compensate for this bias. There is a methodological limitation with this method, however: if the random start distance is a uniformly distributed distance between the specified points then even if the marker is not moved during trials then the mean average of this position will closely resemble the true midway point. Additionally, if settings are very imprecise, it is difficult to measure bias accurately. One way to counter these issues is to ask observers to split the distance into more than two sections. This method of trisection was used in Experiment 1. If far distance is compressed relatively to near space, then it would be expected that both markers would be set at a lesser distance. Conversely, if far distance is expanded, both markers would be positioned further away than the correct values. This is presented in *Figure 4.1*.

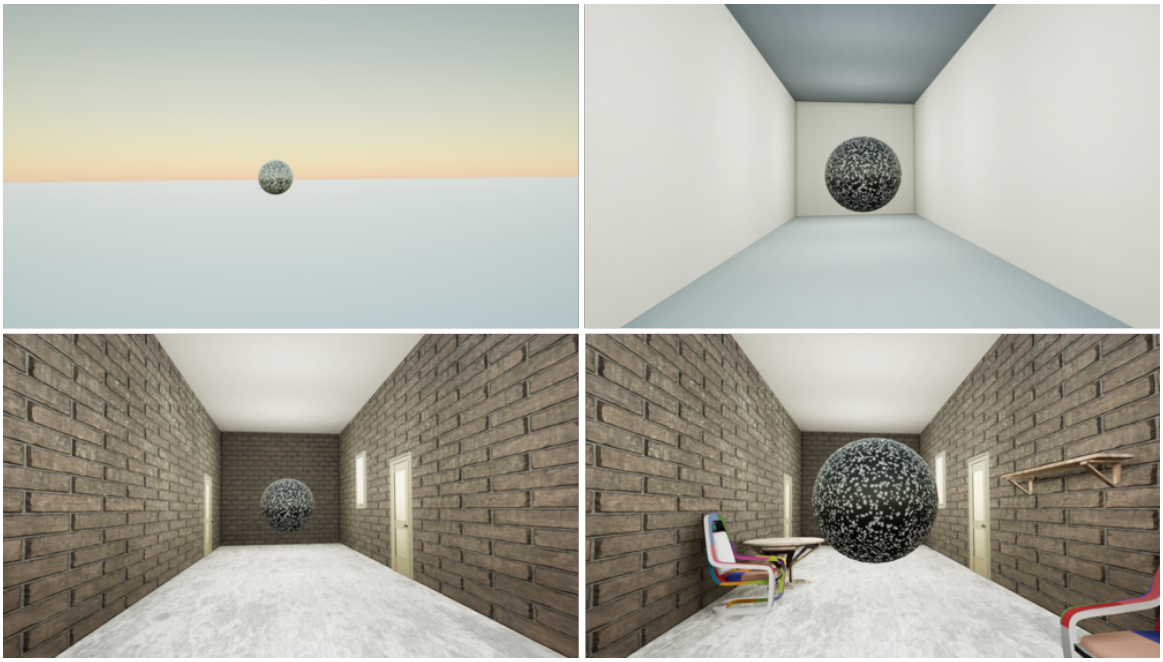


Figure 4.3: The four Constancy task environments: `Sparse` (top left), `Perspective` (top right), `Textured` (bottom left), `Cluttered` (bottom right)

While distance bisection and trisection tasks can be used to measure nonlinear compression and expansion, accurate settings could be made using this method if distances were under- or over-estimated by the same scaling factor at all distances. To assess biases which might also be a linear scaling of distance, a size constancy task (Carlson, 1960; Hornsey & Hibbard, 2018; Kopiske et al., 2019; Murgia, Sharkey et al., 2009; Scarfe & Hibbard, 2006) was used as an indirect measure of the accuracy of distance perception. In this task, observers indicate the size of an object presented at differing distances. If distance is perceived accurately, the indicated size of a stimulus would be accurate. If the size settings are too large, and this error is attributed solely to mis-perception of distance, this would denote the underestimation of distance. Conversely settings that are too small would denote the overestimation of distance. This method was used in *Chapter 3* and found that observers' settings, when viewed in VR, were consistent with a progressively greater underestimation of distance with increasing presentation distance, consistent with previous psychophysical studies using stereoscopic displays (Brenner & van Damme, 1999; Johnston, 1991; Scarfe & Hibbard, 2006).

4.1.3 Current Study

Here, four environments were created, with increasing amounts of pictorial depth cues as shown in *Figures 4.2* and *4.3*. In the simplest condition, the `Sparse` scene merely presented a floor plane, similar to the environment used in *Chapter 3*. In this condition, the principle cues to distance were binocular convergence and the size of the image of each object. Linear perspective was added by presenting walls and a ceiling to create the `Perspective` scene. The `Textured` scene added bricks and a floor pattern in order to include texture as a cue to distance. Finally, in the `Cluttered` scene, object clutter was added to the scene to provide a broad range of pictorial and binocular cues. The step from `Textured` to `Cluttered` added potential anchors into the scene which were able to provide supplemental depth cues through additional disparity information, or the integration of information across distance (Wu et al., 2004). Evidence suggests that texture might be the most important cue for distance perception: Sinai et al. (1999) found that estimates were better when brick floors were added to a stimulus, and Thomas et al. (2002) found depth matching estimates were more accurate with vertical line textures visible. Kline and Witmer (1996) determined that the most accurate perceptions of virtual distances arose from having a wide FoV and a pattern with high resolution. In addition, Loomis, Knapp et al. (2003) hypothesised that any distance compression found in VR might be due to the simplicity of virtual environments, hence the final environment used in the current study incorporated an array of naturalistic objects to the scene.

In the first and second experiment, distance perception was assessed in each of the four scenes using distance trisection and size constancy tasks, respectively. In Experiment 3, repetitions of the `Sparse` and `Cluttered` scenes from Experiment 2 (*Figure 4.3*) were used with both monocular and binocular viewing, to assess the importance of binocular cues to the perception of distance under free viewing. The viewing distances tested were across a range representative of many indoor environments.

Directional biases in the data were observed through the signed errors in each condi-

tion, where the error is the difference between the accurate setting and the participant's setting in each trial. Here, an error below zero would indicate the setting fell short of accurate performance, and errors above zero indicate the participant's setting was greater than accurate. The absolute values of the error (unsigned) were used to identify overall accuracy in responses for each condition.

In addition to measuring systematic bias in the perception of distance and accuracy, the precision of observers' settings was also measured. Precision here was presented in the form of standard deviation of participant responses, however further analyses were also undertaken so that a model of cue combination could be incorporated to the discussion of Experiment 3. For each observer, this was quantified as the variance of settings across trials (σ^2). According to the weighted averaging model of cue combination, perceived distance is an average of the estimates of distance derived from multiple individual cues. Here, how precision is improved by the weighting of pictorial and binocular cues (Landy et al., 2011; Murray & Morgenstern, 2010) can be predicted:

$$\mathtt{CuesCombined} = \mathtt{W}_{\mathtt{pictorial}}\mathtt{D}_{\mathtt{pictorial}} + \mathtt{W}_{\mathtt{binocular}}\mathtt{D}_{\mathtt{binocular}} \tag{4.1}$$

where $\mathtt{W}_{\mathtt{pictorial}}$ and $\mathtt{W}_{\mathtt{binocular}}$ are the weights associated with pictorial and binocular cues, and the distance estimates from these cues are given by $\mathtt{D}_{\mathtt{pictorial}}$ and $\mathtt{D}_{\mathtt{binocular}}$, respectively. These estimates are assumed to be unbiased. Averaging the cues with these relative weights, which are proportional to their inverse variance, will optimally improve the precision (σ) of the resulting estimate ($\mathtt{e}$):

$$\sigma^2_{\mathtt{e}} = \frac{\sigma^2_{\mathtt{binocular}}\sigma^2_{\mathtt{pictorial}}}{\sigma^2_{\mathtt{binocular}} + \sigma^2_{\mathtt{pictorial}}} \tag{4.2}$$

This experiment directly measured the degree to which the precision of distance estimation is improved by the availability of binocular cues. Overall, the aim of this series of experiments was to establish how binocular and pictorial cues contribute to the accuracy and precision of distance judgements in virtual reality, and to assess the contributions of these cues against the prediction of a weighted-averaging model of cue

combination (Landy et al., 2011).

4.2 Experiment 1: Methods

This experiment used a distance trisection task in environments with increasing pictorial depth cues. This allowed us to determine how these cues contribute to the accuracy of distant estimates.

4.2.1 Participants

A total of 40 participants were recruited from the University of Essex's online participant advertisement system (SONA). The ages ranged from 18 to 26 with a mean of 20. Participants were rewarded with course credits, completion time averaged approximately 60 minutes.

4.2.2 Stimuli

The reference stimulus (a statue obtained from Unreal Engine's Starter Content) was used to indicate the distance to be trisected. Two target spheres were set 50 cm either side of the reference stimulus and were referred to as the *far target* and *near target* (both *target stimuli*). These were presented at eye-height.

The other cues in the room were loaded from Unreal Engine Starter Pack, where each environment had specific cues visible:

- `Sparse`: where only a floor surface was visible. Distance information is provided primarily by the size and binocular cues of the reference and target stimuli themselves.
- `Perspective`: with the addition of plain surfaced walls and ceiling. The additional distance information here is provided by the linear perspective from the intersections between the walls, ceiling and floor.

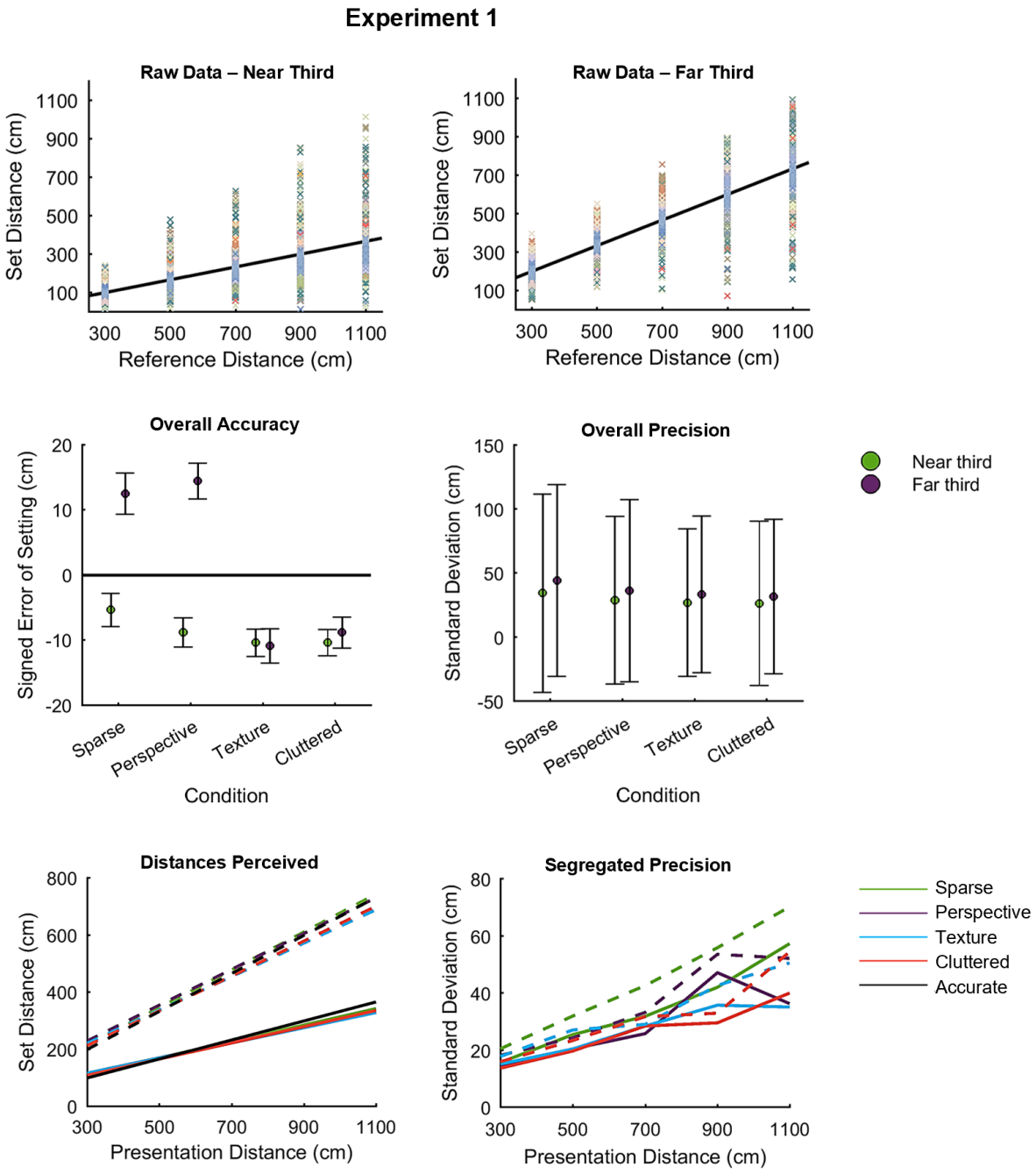


Figure 4.4: Figures from analyses of Experiment 1. Raw data for each participant in a different colour, across all trials. Accuracy and Precision graphs show the near third in green and far third in purple. Distances Perceived and Segregated Precision have the near thirds in solid lines and the far thirds in dashed line for each coloured environment.

- `Textured:` a brick pattern on the wall, and a textured floor covering, provide additional texture cues.
- `Cluttered:` Furniture and shelving created a more cluttered environment, providing a rich, multi-cued environment to provide a natural gamut of depth cues.

The colour scheme and light sourcing were rendered to provide a naturalistic setting, with the same static light illuminating all environments. Each environment is shown in *Figure 4.2*.

4.2.3 Apparatus

A PC with an Nvidia GeForce GTX 1060 graphics card was used to create and present the stimuli, through Unreal Engine 4.18. The Oculus Rift was used to present the environments, with an adult population average inter-ocular distance of 63 mm (Dodgson, 2004) set for all participants. Participants used the Oculus Motion Controller to take part in the experiment. These apparatus were used for all experiments.

4.2.4 Task and Procedure

In each environment, the reference and target spheres were constantly visible. Target stimuli were set to random X axis (forwards/backwards egocentric distance) locations between the participant and the reference stimulus between each trial. Participants had control over the X axis position of the target spheres through button presses on the matching side left/right controller. The aim was for participants to position the near target at one-third of the distance between their-self and the reference stimulus, and the far target at two-thirds of the distance.

After instructions on how to complete the task were given to each participant, they moved a chair into a specified start position which was indicated by a pink circle on the floor in the virtual environment. The reference stimulus was positioned at distances of 3, 5, 7, 9 or 11 metres, varied between trials. Once the participant had decided upon the distances of both target stimuli, they indicated this to the experimenter so the absolute

distances of the targets could be saved and to move onto the next trial.

The order in which the participants completed the environments was counterbalanced, as was the distance of the reference stimulus within them. Each block of trials consisted of 15 repetitions for each of the 5 reference distances (3, 5, 7, 9, and 11 m). These 75 trials were presented in a randomised order.

4.3 Experiment 1: Results

4.3.1 Analysis Strategy

A linear mixed effects model was used to predict the distance settings. For each environment there were two sets of analyses conducted, with either intercept only, or intercept and reference distance ($\mathtt{D}_{\mathtt{reference}}$) as random factors across participant ($\mathtt{P}$). The model that was used predicted the distance setting using the reference distance as a fixed factor and a random intercept and slope and had the lowest AIC out of the two:

$$\mathtt{DistanceSetting} \sim \mathtt{D}_{\mathtt{reference}} + (1 + \mathtt{D}_{\mathtt{reference}} \mid \mathtt{P}) \tag{4.3}$$

Error of participant settings were calculated by subtracting the accurate distance setting of the target from the set distance. This was done in each trial and for the near and far targets separately. A negative result from this indicates the set distance was not as far as the accurate distance; for the unsigned error the absolute value of errors was calculated.

The standard deviation of responses was used as a second measure of performance: in order to normalise the data, the fourth root of the standard deviation were calculated for use in all precision analyses (Hawkins & Wixley, 1986; Todd et al., 2010). The fourth root is the square root of square root of the number, and is used as it has been shown

to produce a nearly symmetric distribution that is very close to normal (Hawkins, 1981).

A one-way ANOVA was use to determine if there was a significant difference in accuracy between the environments. Another one-way ANOVA was used to determine if there was any difference in the precision of estimates between the different environments. Repeated t-tests were then conducted to determine between which environments the significant improvements in performance occurred.

4.3.2 Measures of Performance

The results of the mixed effects model using *Equation 4.3* are shown in *Table 4.1* and plotted in *Figure 4.4*. All near third slopes, in each environment, were significantly different from 0.333 at the p <0.05 level, and for the far third slopes all bar the `Sparse` environment produced values significantly different from 0.666. These are presented in *Figure 4.5*. The intercepts for the near thirds were all significantly larger than zero; all far third intercepts bar the `Sparse` environment were also significantly larger than zero.

The accuracy of distance settings, which are presented as the mean signed error in *Figure 4.4*, show the near third settings were consistently set closer than the true first third distance in all environments, while the far third settings were overshot in the `Sparse` and `Perspective` environments, but also undershot in `Textured` and `Cluttered` conditions. A one-way ANOVA was used with environment as a predictor for unsigned error (near = F (3, 8994) = 6.991, p <0.001; far = F (3, 8994) = 19.615, p <0.001) which led to repeated t-tests being carried out: for near settings a significant difference between `Sparse` and `Perspective` environments (p <0.001) was found; for far settings there was a significant difference between `Sparse` and `Perspective` (p <0.001) and `Textured` and `Cluttered` environments (p <0.001). In both the near and far third settings, there were improvements via a reduced error in estimates when additional cues were incorporated into the environment.

The average of the standard deviation for each participant's repeated settings are presented in *Figure 4.4*. The most noticeable decrease in variation occurred between the `Sparse` and `Perspective` environments. The analysis was a mixed effects linear model, using distance as a predictor and participant as a random structure on both the near and far settings separately, across all four environments. For the near-setting data set, both intercepts and slopes were significantly different from zero at the 0.01 significance level which led to post hoc analyses to compare the improvement of precision in the `Cluttered` environment to all other environments and to further explore the interaction of distance and environment. The improvements from the `Sparse` ($p = 0.014$) and `Perspective` ($p < 0.001$) environments were notable, as was the difference between the `Sparse` ($p = 0.010$) and `Cluttered` ($p < 0.001$). The same analyses were conducted on the far-setting data, revealing the same significant main effect of environment on the precision, and the follow-up regression main effect of condition improvement was significant in the `Sparse` ($p < 0.001$) and `Perspective` ($p = 0.019$); also interactions of condition and distance within `Sparse` ($p = 0.040$) and `Cluttered` ($p < 0.001$) environments.

4.3.3 Summary

This sectioning task required participants to identify the first and second third of total distance between themselves and a stimulus, in four virtual environments. Each environment had different visual cues available, and performance was measured in terms of the distances perceived, the accuracy of estimates, and the precision of responses. Overall, there was less variability in the distances of the near third settings compared to the far, and this variability increased as the presentation distance of stimuli increased further from the participants. Participants were able to complete the task reasonably well, however it is not possible to determine whether participants were able to perceive the absolute distances presented here; it may be that participants were able to determine the relative distances within the environment without understanding what the distances

actually were. For this reason, a constancy task was undertaken in Experiment 2.

4.4 Experiment 2: Methods

In this experiment a size constancy task was used with the same environments assessed in Experiment 1.

4.4.1 Participants

A new group of 20 participants were recruited from the University's online participant recruitment system, who had a mean age of 21 and ranged between 18 and 27.

4.4.2 Stimuli and Apparatus

The same four environments from Experiment 1 were used, but instead of the target and reference stimuli there was only one target sphere. The target axes were locked in that when the size was manipulated either by the participant or between trials, it would increase/decrease as a sphere. A size 5 (22 cm diameter) football was used as a physical reference for participants.

4.4.3 Task and Procedure

Participants held the headset's controller in one hand, and the football in the other/on their lap. Neither were visible to the participant during the trials. Before starting, they moved into the specified start position and were instructed not to move the chair from this position.

In each trial a target stimulus was visible at each of the previously tested egocentric distances of 3, 5, 7, 9 and 11 metres. Controller buttons were used to increase/decrease the size of the target object until the size appeared to match the size of the football participants held. Participants verbally indicated they had completed the trial to the

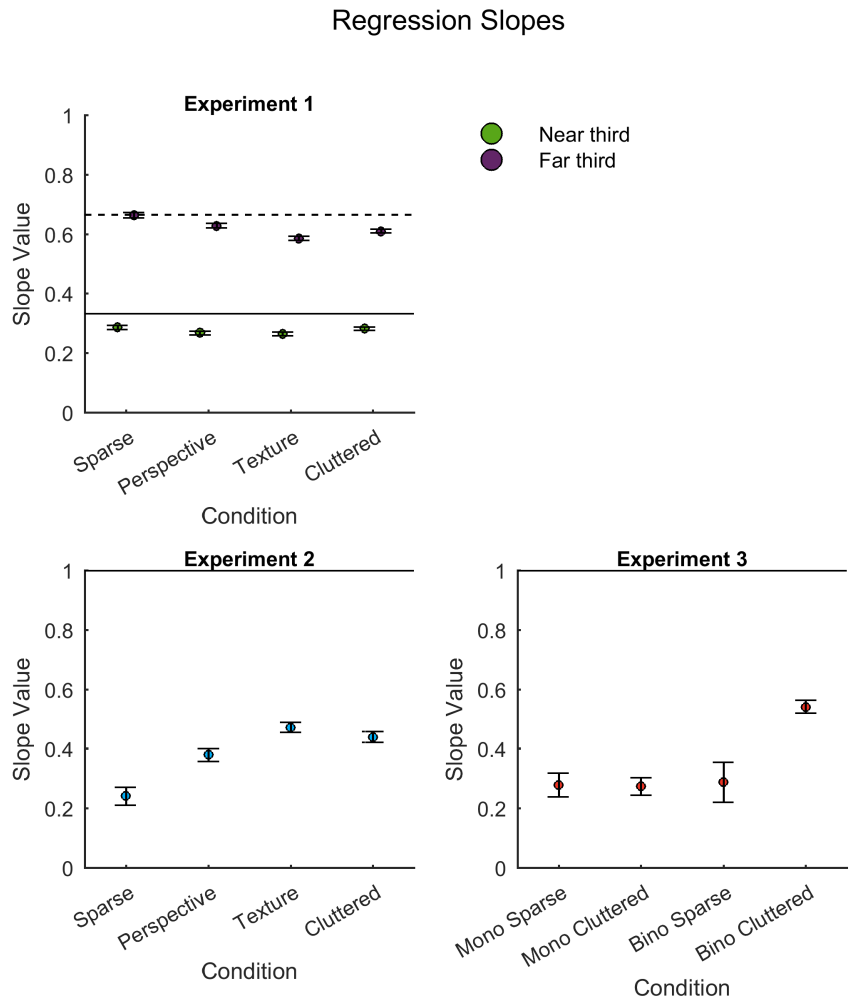


Figure 4.5: Comparison across experiments. Regression slope values for distances perceived in all environments in each experiment, with 95 % confidence limits. Accurate setting indicated by the black lines (Trisection experiment- near third 0.333, far third 0.666 is dashed; Constancy experiments 1).

experimenter, at which point the metrics was saved and moved onto the next trial or new environment.

The order of environments was randomised, as were the distances of the target, and the size of the target in each new trial.

4.5 Experiment 2: Results

4.5.1 Analysis Strategy

The same formatting and analyses from Experiment 1 were performed for the results of this experiment. As this was a constancy task, the error was calculated here by subtracting the correct size of the target sphere from the set size in each trial.

The size settings were used to indirectly measure the perceived distance to the stimuli; the data points were converted to the effective distance perceived in each trial by multiplying the set size of the target in a trial by the presentation distance, then dividing this by the correct size setting.

Another linear mixed effects model was used to analyse these new *perceived distance* data in each environment:

$$\texttt{PerceivedDistance} \sim \texttt{D} + (\texttt{1} + \texttt{D} \mid \texttt{P}) \tag{4.4}$$

This included target distance (D) as a fixed effect and as a random slope; this reduced the AIC compared to a model in which this random slope was not included. Participant was also used as a random effect.

4.5.2 Measures of Performance

Raw data for the size estimates and the effective distances perceived are plotted in *Figure 4.6*. The accuracy data, which shows the decrease in size errors as more environmental cues were made available, and the variance is plotted, which also reduced as pictorial cues were increased. Constancy errors were lowest in the full cue condition, although when converted into the effective distance, the slope value from this and all other environments was nevertheless significantly lower than one (accurate performance), indicating a lack of perfect constancy in size perception across distance.

The results for accuracy, as shown in *Figure 4.6*, were analysed using a one-way ANOVA with condition as a predictor for the unsigned error. The significant output ($F(3, 2994) = 79.172$, $p < 0.001$) led to repeated t-tests to be carried out; significant differences were found between all environments ($p < 0.001$), with the `Sparse` environment presenting the lowest accuracy. Accuracy increased consistently with the addition of the tested visual cues.

The effective distance perceived in each trial are also shown in *Figure 4.6* with accurate performance displayed as the black line. The output values of the mixed effect model of *Equation 4.4* are shown in *Table 4.2*. These show that in each environment, all slopes were significantly different from one (accurate performance) at the $p < 0.05$ level. *Figure 4.5* depicts these slopes for each environment with 95 % confidence intervals. All distances were perceived as being significantly less than what was rendered across each condition, with the most compression found in the `Sparse` environment and least in the `Cluttered` environment.

The standard deviation and 95 % confidence limits are shown in *Figure 4.6*, with the greatest variability found in the `Sparse` environment. The fourth root of the standard deviations were calculated and centred for analysis. Mixed effects models to predict these values (in each environment separately) using distance as a predictor and participant as a random structure revealed significant differences at the 1 % level, showing that the variance in settings were consistently altered by presentation distance across environments. This led to post-hoc analysis of the `distance * environment` interaction and to compare the `Cluttered` environment's improvement against the other environments. No significant interactions were found, although the improvements of precision in the `Cluttered` environment were significantly different from the `Sparse` and `Perspective` environments (both at $p < 0.001$).

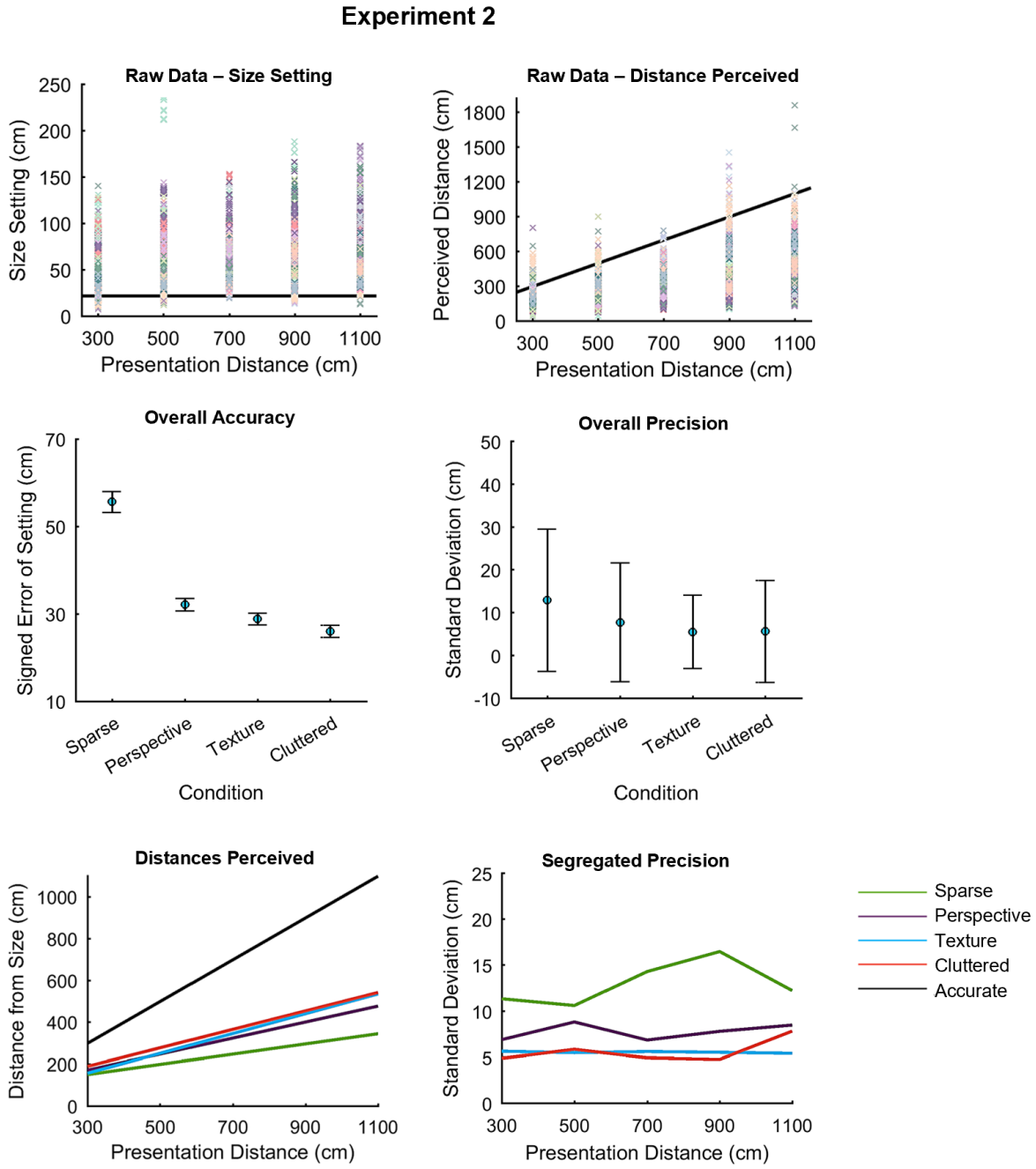


Figure 4.6: Figures from analyses of Experiment 2. Raw data for each participant in a different colour, across all trials. Different environments colour coordinated.

4.5.3 Summary

In this constancy task, participants adjusted the size of a virtual ball they observed in VR in the same four environments which were tested in Experiment 1. The task was to adjust the size of the ball until it matched the size of a physical ball participants held, but could not see. The accuracy of the settings increased when the virtual environments consisted of more visual distance cues, and overall the environments with the fewest observable distance cues resulted in the largest variation in estimates. The distances perceived are shown to be substantially different to the distances perceived in Experiment 1: in the previous experiment it would have been possible for participants to complete the task without an accurate perception of the absolute distances presented to them, as the task required interpretation of relative differences in distance. In the current experiment, the task was an indirect measure of distance perception in which the absolute distances that participants were perceiving was able to be quantified. The impact binocular cues had on perception was not accounted for here, which is why separation of monocular and binocular cues was tested in Experiment 3.

4.6 Experiment 3: Methods

This was a replication of Experiment 2, using only the `Sparse` and `Cluttered` environments, however each participant viewed the two scenes both monocularly and binocularly. This allowed us to directly test the contribution of binocular cues to improving the accuracy and precision of distance perception in VR across a range of distances.

4.6.1 Participants

Another 20 participants were recruited from the University's advertisement system, who ranged between 18 and 26 and had a mean age of 20.

4.6.2 Stimuli and Apparatus

The `Sparse:` and the `Cluttered` environments from Experiment 2 were used. The conditions were:

- `Monocular Sparse:` viewing through one eye, only a floor surface was visible
- `Monocular Cluttered:` naturalistic corridor components through one eye
- `Binocular Sparse:` the floor surface through both eyes
- `Binocular Cluttered:` corridor components through both eyes

The same size 5 football was used as a physical reference for the task. When in the monocular conditions, tissue was used to cover the left lens of the headset.

4.6.3 Task and Procedure

The same start position was used for participants; holding both the controller and the football in each hand. The two face buttons on the controller increased/decreased the size of the virtual sphere and when participants had matched the size of this to the size of the physical football the experimenter manually saved the measurements and moved the experiment onto the next trial.

The virtual sphere was presented at 3, 5, 7, 9, 11 m in front of the participants. The start size of the virtual sphere was randomised between each trial, the order of presentation distance was also randomised, and the condition order was counterbalanced.

4.7 Experiment 3: Results

4.7.1 Analysis Strategy

The analyses used for measuring performance in Experiment 1 and 2 were once again used for this experiment. Error in size estimates for each trial were calculated, and the variation of responses was determined via standard deviation. A series of ANOVAs

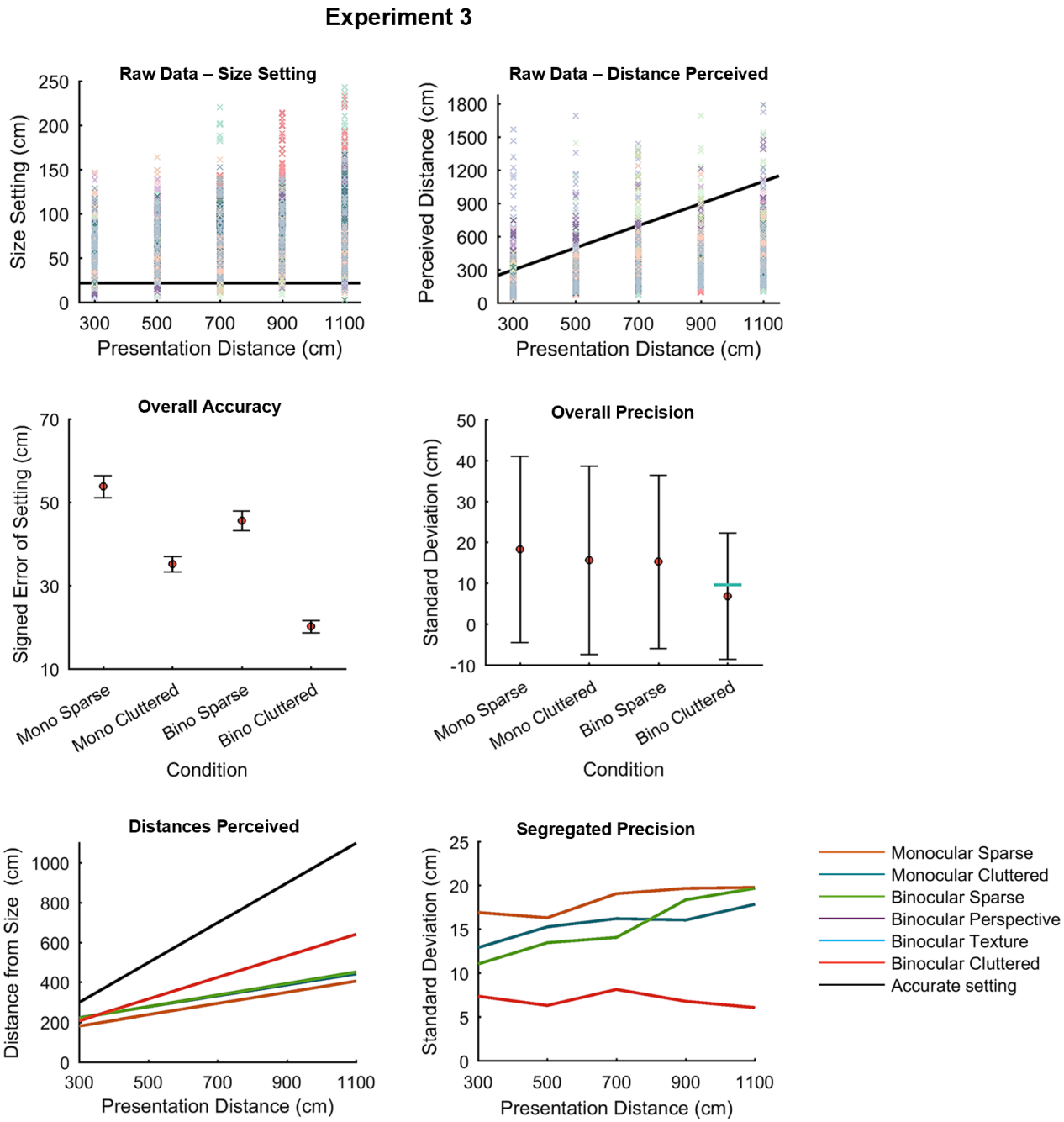


Figure 4.7: Figures from analyses of Experiment 3. Raw data for each participant in a different colour, across all trials. Different environments colour coordinated. Variation prediction of cue combination model in mint.

were used to identify and significant effects between conditions and t-test were then used to on specific conditions from the results of the ANOVA. For the mixed effects model, *Equation 4.4* was used here. The predicted variability in responses, according to the weighted model of cue combination, was calculated based on the fourth root of the standard deviation of responses from two of the conditions tested; this was compared to results obtained from the most comprehensive condition.

4.7.2 Measures of Performance

The raw data from this experiment is presented in *Figure 4.7*. Looking at performance, adding environmental cues (clutter) to the `Monocular Sparse` condition reduced the mean size errors by 18.66 cm, and adding binocular disparity to the sparse environment decreased size errors by 8.24 cm. The addition of these two cues also influenced the precision of participants' size estimates, whereby the amount of variability in the `Binocular Cluttered` environment is significantly (15.09 cm) less than that in `Monocular Sparse`.

The same regression analyses ran in Experiment 1 and 2 were conducted on the perceived distances for each environment, again with the best AIC resulting from including reference distance as a random slope (*Equation 4.4*). In *Table 4.3* it can be seen that only the intercept value for the `Binocular Cluttered` environment is *not* significantly different from zero, showing most accurate performance in this condition. Slopes in all conditions are significantly lower than one, as plotted in *Figure 4.5*, showing again incomplete constancy, but best performance once again in the `Binocular Cluttered` environment.

A one-way ANOVA using environment to predict the amount of unsigned error produced a significant effect: $F\ (3,\ 2994) = 78.409$, $p < 0.001$. Follow-up t-tests revealed significant differences between all environments at $p < 0.001$, showing that the `Binocular Cluttered` environment achieved the highest level of accuracy.

After centring the data, a significant (at the 0.01 significance level) categorical mixed

effects model was conducted to compare the fourth root of the standard deviations of each environment to that of the Cluttered results. This was to check if the change in precision occurred only at some distances for any environment and revealed all interactions were significant, with `Monocular Sparse` at $p = 0.050$, `Monocular Cluttered` at $p = 0.013$, and `Binocular Sparse` at $p < 0.001$. The intercept improvement from all environments to the `Binocular Cluttered` were also significant at the 0.01 % level.

A final t-test was conducted to compare the predicted precision of the full cue condition to the actual results. The fourth root of the standard deviations for each environment were used to obtain the predicted standard deviation, which was calculated as:

$$\sigma_{\mathtt{e}} = \frac{\sigma_{\mathtt{MonocularCluttered}}\sigma_{\mathtt{BinocularSparse}}}{\sigma_{\mathtt{MonocularCluttered}} + \sigma_{\mathtt{BinocularSparse}}} \tag{4.5}$$

where σ_e is predicted standard deviation in the full cue condition, $\sigma_{\mathtt{MonocularCluttered}}$ is the standard deviation in the `Monocular Cluttered` condition, and $\sigma_{\mathtt{BinocularSparse}}$ is the standard deviation in the `Binocular Sparse` condition. The reduction in standard deviation was significantly greater than predicted, T (99) = -5.658, $p < 0.001$: precision was better than expected in the full cue environment tested here.

4.7.3 Summary

This constancy task replicated two of the environments from Experiment 2, to separate the impact both pictorial and binocular cues had on distance perception in VR. Here, distances were generally perceived as being around half of what was presented; the most amount of variability and errors were produced in the environment where minimal pictorial cues were present, the absence of binocular cues and scene clutter caused participants to perform the worst. Performance was highest in the environment where an abundance of binocular and pictorial cues were available. The findings here were consistent with the weighted model of cue combination, for the pictorial and binocular

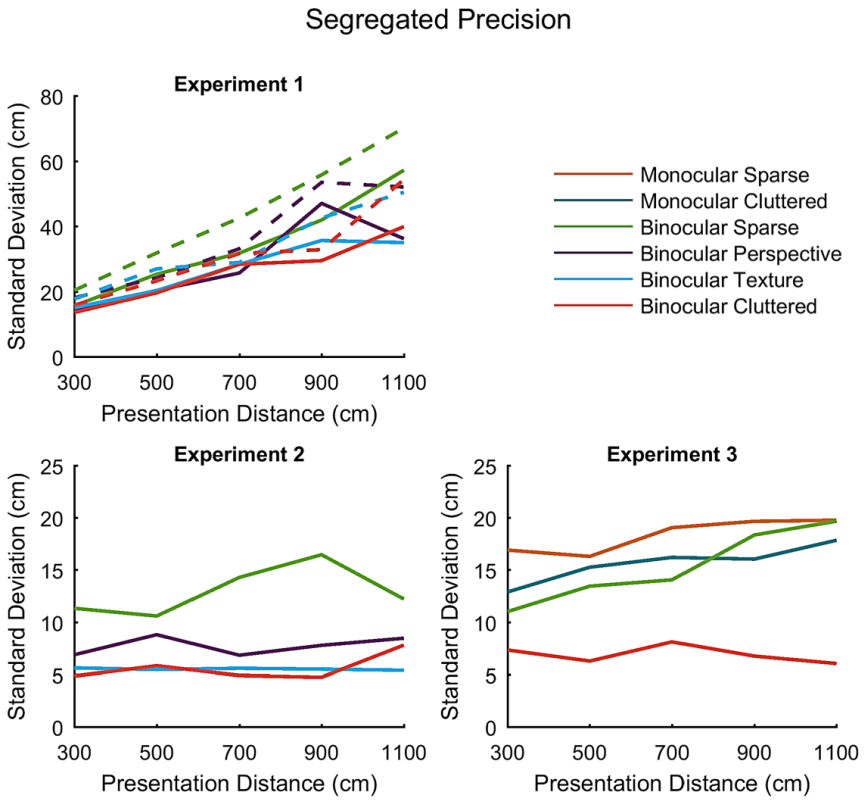


Figure 4.8: Comparison across experiments. Overall trends of standard deviation for each environment in all experiments. In Experiment 1 data for near thirds indicated by solid lines, far thirds are dashed.

cues tested here.

4.8 Discussion

The effects of the availability of pictorial and binocular cues on the accuracy of distance perception in VR were assessed using trisection and size constancy tasks. The trisection experiment was a direct measure of distance perception, whereas the constancy experiments were indirect. The bias and precision of distance perception using these tasks were assessed.

4.8.1 Bias and Precision

Each of the experiments tested the effect specific visual cues had on performance, which was measured in terms of accuracy and precision. In relation to this however, there is a distinction between precision and bias that is not always taken into consideration when using cue combination models. Through the analysis procedure used here, a bias of (linear) underestimation has been revealed: in *Figure 4.5* the regression values all less than one indicate this. As well as this, *Figure 4.8* shows the standard deviation varying across distances and environments which instead demonstrate the metric of precision.

In studies of cue combination, there is the assumption that there will be no bias arising from either of the two factors being tested (here: distance, environment) along with the assumption that the factors will create random variation (Scarfe & Hibbard, 2011). In this instance however, it has previously been established that distance perception is biased in the direction of underestimation. However, it was still predicted that precision would improve if cues can be combined through appropriate weighted averaging.

In Experiment 3, the precision in the full cue condition can be predicted from the data in reduced cue conditions using *Equation 4.2*: the contribution of binocular disparity *and* pictorial cues were able to be predicted for the `Binocular Cluttered` environment. By using the values from *Figure 4.9* (`Monocular Cluttered` = 15.682; `Binocular Sparse` = 15.343), *Equation 4.5* gives the predicted standard deviation of 9.711, CLs [-4.192 23.615] if the two sets of cues were combined optimally. The results give an actual standard deviation of 6.934, CLs [-8.486 22.354], which are consistent with the predicted improvements from cue combination, as this result is within the credible range of the wide estimated standard deviation. The 95 % confidence limits for the fourth root of the standard deviation and prediction however, do not encompass one another.

The findings of the current study are consistent with prior research into the contribution of stereoscopic viewing: the binocular cues added when viewing a scene stereo-

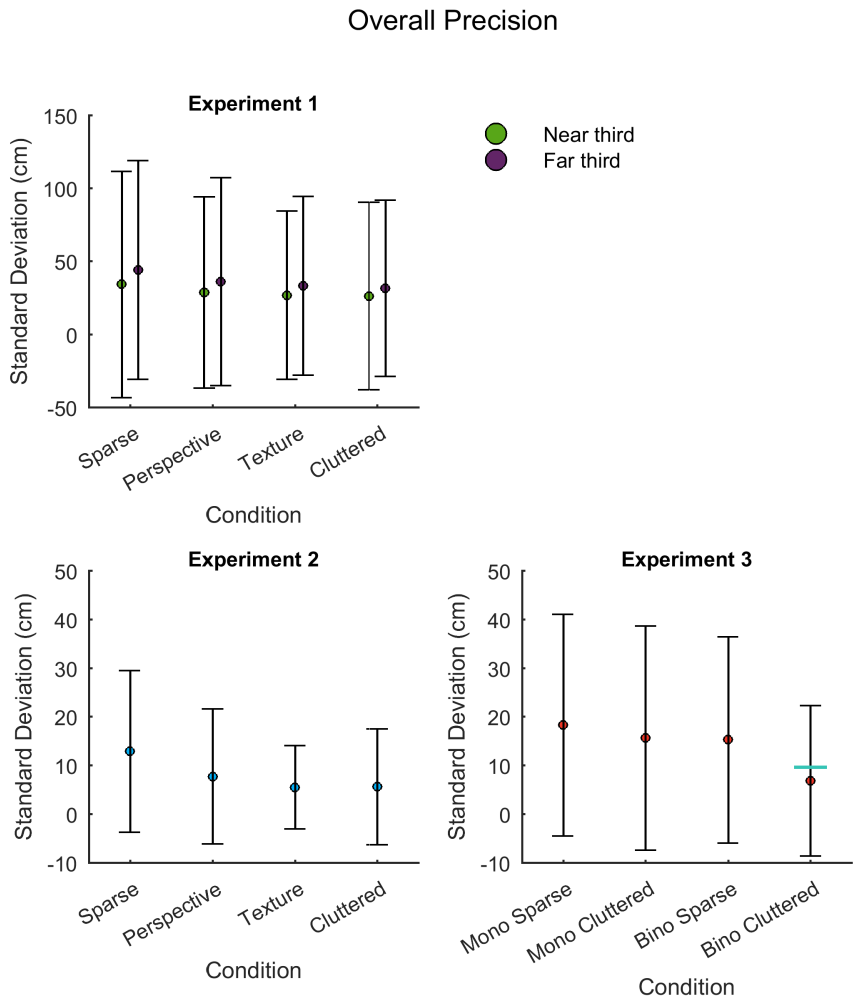


Figure 4.9: Comparison across experiments. The amount of variation in responses calculated from the average standard deviation for each participant at each distance, plotted for each environment. Ideal performance would have zero variance, 95 % confidence limits as error bars. In the Trisection task first showing the near third (green) then far third (purple) setting. Experiment 3 shows variation prediction of cue combination model in mint.

scopically can improve sensitivity to depth (Hornsey et al., 2015). It has been concluded that binocular disparity provides complimentary depth information to the monocular pictorial cues. Just like the findings in the current study, when combining the two types of cues the performance is significantly enhanced to be greater than any one subset of cues alone.

An alternative model of cue combination is the cue veto model (Bülthoff & Mallot, 1988), which suggests unreliable cues are ignored rather than having any weighting attributed to them. While there may be some evidence for that model with regards to

other cues (Hartle & Wilcox, 2021), the weighted averaging model presented here appropriately accounts for the combined effect of both monocular and binocular cues.

A study by Kim et al. (1987) compared performance in a task using a monocular and a stereoscopic display: it was found that incorporating a texture grid with ground intercepts produced performance that was equal to that of the stereoscopic display alone. This is similar to the findings of Experiment 3 (*Figure 4.10*), that the distances perceived from monocular viewing with full pictorial cues were uniform with distances perceived binocularly but with no pictorial cues. Using this finding for future work with similar equipment, or wanting to get the same desired outputs; developers should only use cues which aid 3D vision in stereo-displays as it will reduce complexity of computer system. If costly cues like binocular disparity are just as effective as simple pictorial cues then the least costly option should be chosen.

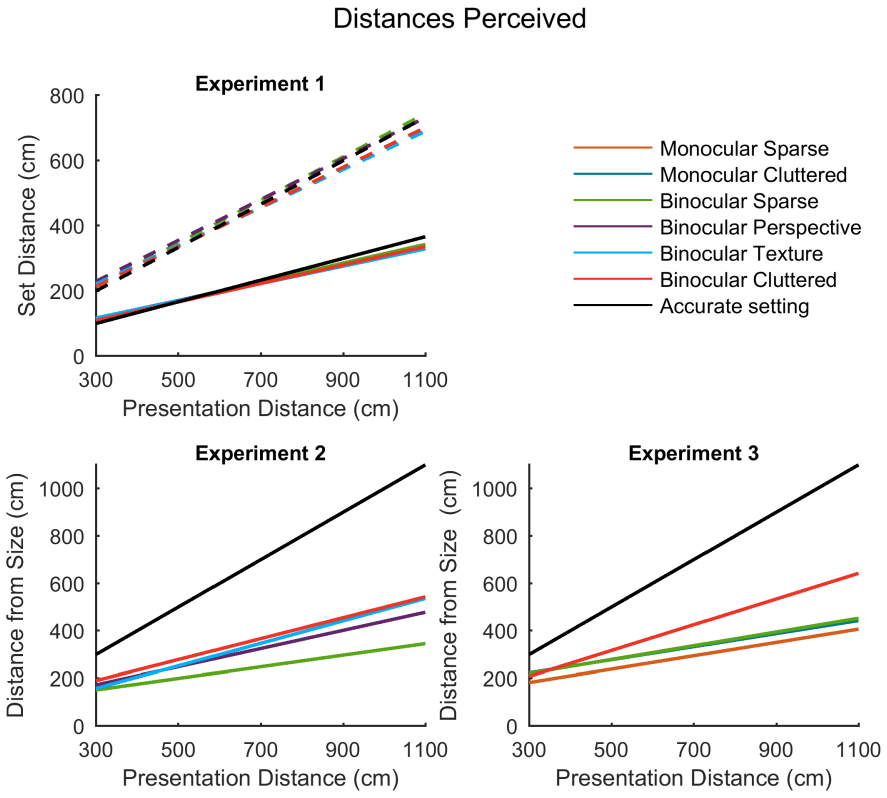


Figure 4.10: Comparison across experiments. Overall regression trends for the distances perceived in each experiment. Accurate performance indicated in black. Experiment 1 values for far third indicated by dashed lines, near third are solid lines.

4.8.2 Comparing to the Hypotheses

The first hypothesis, that there would be an underestimation of near space compared to far, in Experiment 1 could be shown by either: the near third position consistently set too close but with an accurate far third; or setting both thirds too close when the reference stimulus is at the closer distances but accurate setting at the further distances tested. This systematic underestimation of the near third (of around 8 cm) is found in the results of *Figure 4.11* pointing to the underestimation of near space relative to far. In Experiments 2 and 3 this would be portrayed by size settings becoming larger as the presentation distance increases. *Figure 4.12* indeed shows that overall, there was an underestimation of near space as an increasing amount of the size settings are above the black line of accurate performance as presentation distance becomes greater. Following this up with statistical analyses on the impact of presentation distance on perceived distance (derived from size set): it can be seen in *Figure 4.5* that in the constancy tasks, the effect of presentation distance was reduced when visual cues were added, as the slope gradient for each environment's regression approaches one (whereby perceived distance is equal to presentation distance).

The second hypothesis, that performance would improve with the addition of cues, was also supported by the results from the constancy experiments. In *Table* 4.2, the slope values increase with each cue that is added (accurate performance slope = 1). The values are still less than one, indicating incomplete constancy. Accuracy, measured by signed error, is shown in *Figure 4.11* to improve with each additional cue in Experiment 2, and the addition of context and disparity separately in Experiment 3. Experiment 1 only showed improvement of accuracy from the addition of visual cues from `Sparse` to `Perspective`.

The other aspect of performance was measured by the standard deviation of settings across participants in each condition and distance. In *Figure 4.9*, each graph shows how the variability in participants' performance differs with the addition of cues. The overall

standard deviation values decrease with the addition of cues which is a result of a higher sensitivity to distance in each trial. Taking the results for both accuracy and variability into consideration, the conclusion drawn is that performance is indeed enhanced with the addition of these specific visual cues.

Experiment 3 replicated the `Sparse` and `Cluttered` environments with binocular and monocular viewing, to determine the contribution of binocular disparity compared to pictorial cues. It was hypothesised that higher performance would be obtained by binocular than monocular viewing in the `Sparse` condition; and that the `Cluttered` environments would increase performance than the `Sparse` environments. The t-test analyses found evidence for binocular viewing significantly improving the perceived distances, seen in *Figure 4.10*. The second part of this hypothesis was that there would be less of a difference found between the `Monocular Cluttered` and `Binocular Sparse` conditions, which was indeed found to be a non-significant change in both the precision and accuracy tests.

4.8.3 Interpretation of Findings

Performance was worst in the `Sparse` environments in each experiment, shown by the combination of inaccurate slope and intercepts in *Tables 4.1, 4.2,* and *4.3*. Murgia, Sharkey et al. (2009) also found that distances were underestimated in both poor and rich cue conditions, but there was greater underestimation in the poor cue environment. Here, the introduction of more pictorial cues aids participants' perception both in the form of precision and accuracy of settings. In Experiment 1, the addition of each positively impacted the precision, and in Experiment 3 the addition of context and disparity had an extremely positive impact on both accuracy and precision. However, Surdick et al. (1997) suggested that linear perspective and texture gradient are among the most useful cues for distance perception. In Experiment 2, the set size of the target is most level across distances in the `Textured` environment (slope value nearest one in *Figure*

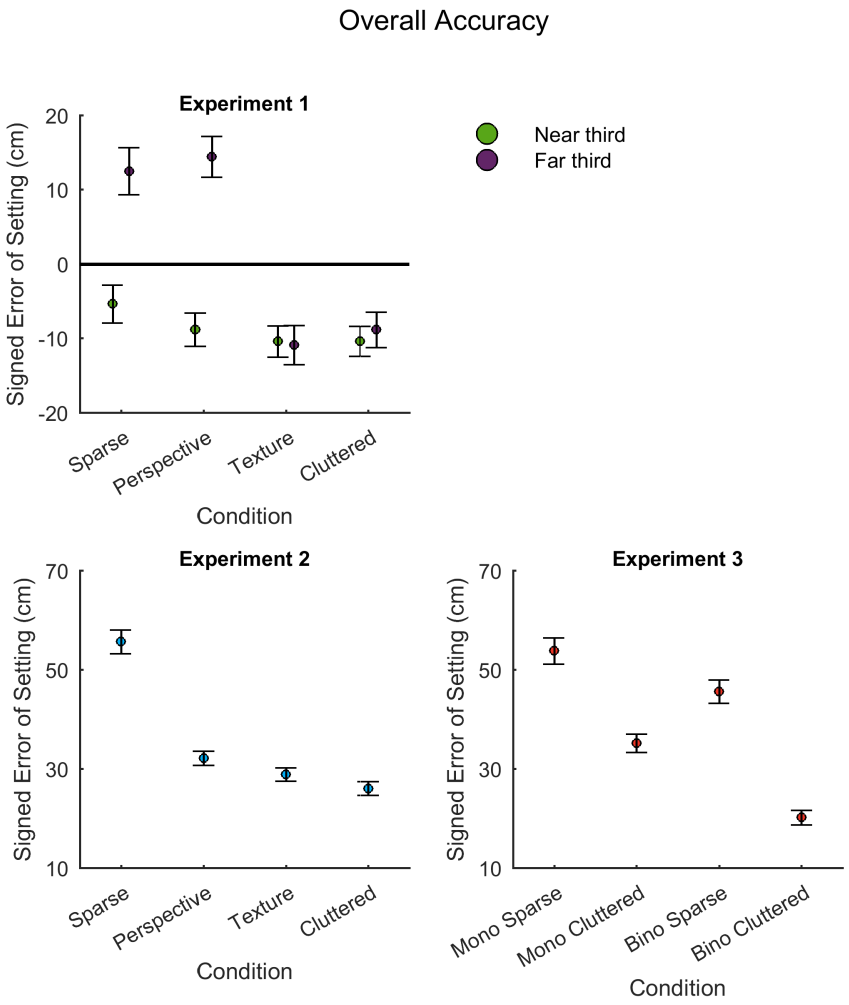


Figure 4.11: Comparison across experiments. Mean signed error for all environments in each experiment, with 95 % confidence limits. Accurate performance would have zero mean error, negative values indicate under-setting of participant response.

4.5), however the confidence intervals for the slope value of the `Textured` environment do not include one, indicating significantly under-constancy. The low intercept of 16 cm obtained from this environment in *Table 4.2* suggests that texture is the best cue for distance estimation at the very close distances, although the slope value is still just under half of what it should be, meaning the stimuli here are still mis-perceived. From these differing results, it could be concluded that the addition of these specific cues is task-dependent for assisting in minimising response accuracy, but there is great evidence that this has a positive effect on minimising response variability.

In addition to this, the findings of the current study support those of Livingston et al. (2009), who investigated whether the perspective cues found indoors, such as the

linear alignment of the floor-walls-ceiling, improve outdoor distance perception when using augmented reality. An underestimation of distance was found in the indoor environment, much similar to the findings here. In general, perspective cues require a ground intercept, which is why the baseline condition presented this initially as opposed to the stimuli floating in space. From this, each additional cue was easily added to the previous environment.

The headset used here has a fixed accommodation distance, unlike in physical space where accommodation is a distance cue that changes according to the fixation distance. A set-up by Bingham et al. (2001) found overestimates in a limited cue experiment, and when the focal plane was reduced by 2 diopters, the overestimation was halved. One might suggest this factor could be influential over the perception of the surroundings in these environments, however research has suggested that accommodation does not have much of an impact when in full-cue conditions (Mon-Williams & Tresilian, 1999, 2000). In addition, the cue of accommodation becomes less reliable with age (Ellis & Menges, 1998), so given the spread of ages used in the three experiments here, the constant accommodative distance is less likely to have an impact as other factors. There are displays available which allow for a varied focal distance, for example Watt et al. (2005b) created and tested a multi-focal device, which showed that any issues surrounding a fixed accommodation distance can be alleviated with such equipment.

It could be argued that the indoor environments used here might limit the applicability of the findings: Lappin et al. (2006) collected depth judgements from open outdoor environments to find that they were nearly veridical, but indoor environments produced an expansion of perceived distances and more variability in responses. However, Bodenheimer et al. (2007) expanded on this with different distances (15 and 30 m), with compression being perceived only at the far distance, and found no significant difference for indoor or outdoors. Future research replicating the current methodology and equipment in an outdoor setting would be useful in identifying if there is a difference caused by having walls and a ceiling indoor, which emphasise the cue of linear perspective.

Another argument could be made about the cue of motion during Experiment 1 when participants adjusted the distance of the target stimuli. Motion parallax has been shown to be a cue that is used during distance estimation (Bishop, 1989; Young et al., 1993), and so given that the movement speed of targets was constant, how quickly the targets appeared to move at a given position in the environment could provide information as to how far they are from participants. This additional cue is not present in the other two experiments, which could be the reason why performance was so high in Experiment 1. However, the post hoc analysis from a previous experiment (see *Section 3.3.2*) revealed no significant improvement in participants who utilised the cue of motion parallax when making their judgements.

Although there was improvement of estimation with additional cues, one could also say that the improvement was not dramatic when compared to performance in the baseline (sparse) conditions. That is, participants were able to estimate distance reasonably well without the additional tested cues. In both Experiment 1 and Experiment 2, binocular disparity and vergencd were available in the `Sparse` environments, which could explain why performance was not simply random noise. This is why Experiment 3 manipulated these cues by including monocular viewing to remove these binocular cues. Perception of distance from all pictorial cues was equal to the perception of distance from binocular disparity alone (Experiment 3, *Figure 4.10*). This is consistent with findings of certain studies (Creem-Regehr et al., 2005; Eggleston et al., 1996; Roumes et al., 2001); however there is evidence for the reducing effect of binocular disparity with increase of presentation distance based on *Chapter 3*. The effect of distance on the effectiveness of pictorial cues however remains unclear. Further investigation into this would be useful to identify the effects of both at even further distances, and to test if the relationship between them remains equal.

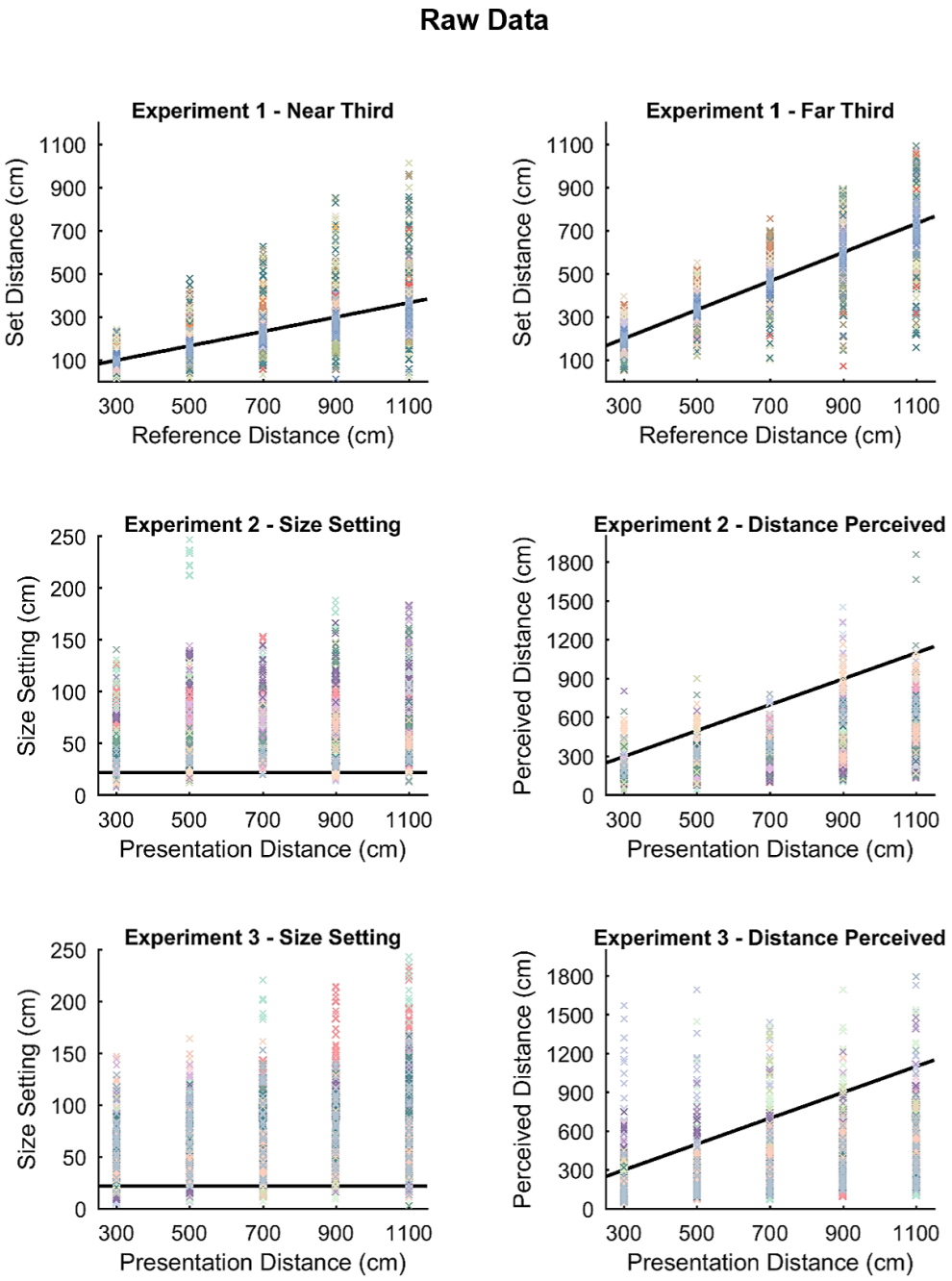


Figure 4.12: Comparison across experiments. Raw data for each participant in a different colour, across all trials. Each row for each experiment with accurate setting in black. One data point was excluded from the final graph for visual consistency with other graphs, but was included in the analysis.

4.8.4 Possible Explanations of Findings

The accuracy results of Experiment 1 show relatively small errors (average of 10 cm in either direction) but this does not necessarily mean that participants were perceiving the rendered distances correctly. Apparent size of objects decreases as a function of distance between it and observer: at far distances changes in apparent distance are less apparent than at closer distances (Sedgwick, 1986). The near and far target spheres in the current experiment were the same size meaning that participants could have compared the size so that the near would be double the size of the far target in the retinal image. This could be an explanation for the results, however even by doing this, participants would still need to set the distance of one of the targets correctly. Because of this, a more plausible explanation is that the relative distances between the stimuli were being perceived accurately. In other words, an observer may section the 9 m between their-self and the reference stimuli equally by positioning the near target stimuli at 3 m and the far at 6 m, but they might be perceiving the total distance as 30 m and positioning the target stimuli at 10 and 20 m. A way to test this would be to replicate the same study, but have the near and far targets appear consecutively as opposed to simultaneously.

In both the `Sparse` and `Perspective` environments of Experiment 1, the far thirds were set further away than where they should have, at about the same magnitude of the under-setting of the near thirds. This is indicative of near space expansion and far space compression. Interestingly, this shifts to an overall compression of space in the `Textured` and `Cluttered` environments. A possible explanation for the perception of half-distance could be that an overload of visual information of familiar objects with accurate physical properties causes confusion for the sizing of the floating target stimulus. It is most apparent in the `Cluttered` environment that the target object does not *belong* in the environment like the other objects, so this might impact on the ability to judge its distance relative to the context. It could be argued that this is the most realistic out of all the environments so that is the reason why performance is worse there when the

target does not fit in: Singh et al. (2010) found that presence of a highly salient physical surface does have an effect on distance perception, but not in a systematic way. This switch to a generally consistent, yet inaccurate perception of the space in the `Cluttered` environment, from a distorted perception in the `Sparse` ones could be a clue as to the specific cues which are favoured by the visual system.

An alternate explanation for the mis-perception of distances seen in *Figure 4.10* is that at specific distances there may be certain cues which are ignored. A study by Glennerster et al. (2006) found that observers ignored available stereo and motion cues when making distance estimates in an immersive HMD, and relied on differing cues more heavily at differing distances. It was found that as viewing distance increased from 1.5 to 3 to 6 m, observers used information from the surrounding clutter objects (such as relative size) as anchors for their estimates, rather than the information from binocular disparity. These findings were taken from a set-up with a much more limited distance range than what has been presented here, however the explanation may be relevant: at further distances binocular disparity would be ignored and pictorial cues more heavily relied upon; this would be represented by more variable responses at further distances in both binocular environments, along with more stability in the estimates of the monocular environments. In *Figure 4.8*, the trend in precision over all distances suggests the most precise estimates are at closer distances, compared to far. The effects or pictorial and binocular cues are approximately summative at 3 m during Experiment 3, in that the variation shown in `Binocular Cluttered` is the sum of both these conditions separately: they are equally as effective at enhancing perception. At around 9 m they diverge and then at 11 m the improvement is being driven almost entirely by monocular cues. This could be argued to be coinciding with the non-linear compression prediction in *Figure 4.1*, however the trends within the raw data plot of *Figure 4.12* in fact present a linear relationship between perceived and presented distance. Perhaps the reason the non-linear relationship found by Glennerster et al. (2006) was not found here was due to the difference in paradigm. Their method intended to instigate a cue conflict between pictorial and binocular information, whereas that was not done here.

A potential cause of the degree of variability in responses throughout each of the three experiments might be from the distance of the two images within the headset being kept the same for every participant, despite participants not all having the same distance between their eyes. Drascic and Milgram (1996) concluded that even small alterations in interpupillary-distance can cause large distortions. Future research may want to record participant's interpupillary-distance, not necessarily to change the headset, but to identify any relationship between this and performance. An additional explanation for the amount of variability across trials could be due to the virtual spheres not being attached to the floor: when an object floats above the bottom plane, the relative height and distance from the observer becomes more difficult to interpret (Drascic, 1991; Kim et al., 1987). For this reason, the tasks used here may have produced results which may be more variable than if the stimuli were attached to the floor plane, much like what would be observed in physical settings.

Some participants may have understood the task straight after viewing the first environment, however others may have needed extra time to fully get to grips with the task at hand: Drascic (1991) tested stereoscopic and monoscopic displays to find that stereoscopic displays enhanced initial learning of the task. This performance advantage existed even after large amounts of practice for some participants. If this were the case in Experiment 3, those who were better-able to understand the task from doing a binocular condition first, compared to half of the participants who did the monocular condition first, would have been filtered out, as the order of conditions were counterbalanced.

Renner et al. (2013) concluded that in order to enable the best possible perception of distances, quality graphics, binocular disparity, a strong sense of presence, and a texture-rich environment are essential. In the current study, it can be argued that binocular disparity and a texture-rich environment were the only aspects which were directly addressed. Presence was not measured, and the graphics of the headsets were not nearly as high as the capabilities that Ultra-HD screens and advanced CGI allow for. Perhaps, increasing these two features may lead to a higher degree of accuracy: an

experiment using the same procedure but with more advanced technology with better rendering capabilities, and also a measure of presence, may produce more accurate results. Additionally, the lighting within the scene may have been an additional indicator to participants about the distances within the scenes. Although the lighting was kept consistent throughout the different environments, the actual position of the light sources (although not visible to the observer) may have been used, as the relative brightness of object decreases as a square of the distance between this object and light source (Surdick et al., 1997). The specific effect lighting has on distance perception is an area to be explored further in VR.

4.8.5 Conclusion and Impact

Overall, the results show systematic underestimation of distance, as well as the accuracy and precision of distance perception improved with the addition of both pictorial and binocular cues. For pictorial cues, precision and accuracy both improved most with the addition of linear perspective in the scene. Binocular disparity and pictorial cues were both found to enhance distance perception. The benefits of binocular disparity, in particular, are important to note. Bias and variability were both reduced by the presence of binocular cues, across a distance range typical of many indoor environments. This improvement illustrates the importance of binocular cues, given the costs of the additional resources required to provide binocular information, and the additional potential for viewing discomfort that this creates (Banks et al., 2012; O'Hare et al., 2013; Shibata et al., 2011b).

Table 4.1: Regression Analyses of each environment in Experiment 1 from Equation 4.3. Accurate performance for the near third would have a slope of 0.333; accurate performance for the far third would have a slope of 0.666. Accurate intercepts would have a value of zero.

Environment		**Intercept Estimate**	**Intercept CLs**	**Slope Estimate**	**Slope CLs**
Sparse	Near	26.390	[15.195 37.585]	0.288	[0.267 0.309]
	Far	14.385	[-7.541 36.311]	0.664	[0.626 0.702]
Perspective	Near	35.898	[24.260 47.535]	0.270	[0.249 0.291]
	Far	41.265	[25.233 57.297]	0.628	[0.604 0.653]
Texture	Near	36.699	[26.199 47.199]	0.266	[0.245 0.287]
	Far	45.202	[30.501 59.904]	0.587	[0.560 0.613]
Cluttered	Near	24.365	[14.683 34.046]	0.284	[0.267 0.300]
	Far	30.461	[15.549 45.373]	0.611	[0.590 0.631]

Table 4.2: Regression Analyses of the distance perceived in each environment of Experiment 2, from Equation 4.4. Accurate performance would have an intercept of zero and a slope of one.

Environment	**Intercept Estimate**	**Intercept CLs**	**Slope Estimate**	**Slope CLs**
Sparse	78.998	[19.030 138.965]	0.243	[0.149 0.337]
Perspective	58.379	[33.428 83.331]	0.382	[0.297 0.466]
Texture	16.414	[-9.071 41.899]	0.473	[0.395 0.551]
Cluttered	58.881	[26.885 90.877]	0.440	[0.373 0.508]

Table 4.3: Regression Analyses of the distance perceived in each environment of Experiment 3 using Equation 4.4. Accurate performance would have an intercept of zero and a slope of one.

Environment	**Intercept Estimate**	**Intercept CLs**	**Slope Estimate**	**Slope CLs**
Monocular Sparse	99.079	[27.181 170.978]	0.280	[0.147 0.414]
Monocular Cluttered	140.989	[69.611 212.368]	0.275	[0.199 0.351]
Binocular Sparse	134.775	[6.613 262.938]	0.289	[0.148 0.431]
Binocular Cluttered	46.370	[-11.573 104.314]	0.542	[0.449 0.635]

Chapter 5

The Role of Cast Shadows in the Perception of Distance

Shadows in physical space are copious yet the impact of specific shadow placement and their abundance is yet to be determined in virtual environments. The previous two chapters (*Chapter 3* and *Chapter 4*) explored the role of projective cues, such as binocular vergence and disparity, and linear perspective, on the scaling of images for distance in the perception of shape and size. The current chapter focuses on lighting cues, in particular the role that shadows can play in the perception of distance. Six lighting conditions were created and presented in virtual reality for participants to perform a perceptual matching task. The task was repeated in a cluttered and sparse environment, where the number of cast shadows (and their placement) varied. This experiment aimed to identify whether a target's shadow was used as a distance indicator in favour of binocular disparity, which was measured by the directional bias of distance estimates.

5.1 Introduction

In virtual environments, just as in the real, physical world, both projective and lighting cues provide information about three-dimensional structure (Todd, 2020). The presented research in this thesis so far has focussed on the role of projective cues, such as binocular vergence and disparity, and linear perspective, on the scaling of images for distance in the perception of shape and size. The current study focusses on lighting cues, in particular the role that shadows can play in the perception of distance. When creating a virtual environment, lighting is required to allow stimuli to be perceived by a user. When rendering environments in software such as Unreal Engine or Unity, both directional and diffuse lighting can be imported, and variables such as the intensity and colour of the light can be manipulated. Individual decisions on the nature and positioning of lighting will all contribute to the quality of information about distance that is available to the viewer. In this study, the effects of a variety of lighting techniques on the perception of distances within virtual reality (VR) were assessed. These experiments were designed to provide a clearer understanding of the contribution of shadows to the perception of distance, and how this interacts with the contents and layout of the scene.

5.1.1 Environmental Context

Many visual cues to distance are available in natural environments (Cutting & Vishton, 1995), and the precision and accuracy with which judgements can be made will depend on the particular combination of cues available in a given context (Hibbard, in press). For example, Surdick et al. (1997) found that in a stereoscopic display, perspective cues were more effective at portraying an accurate perception of distance than other depth cues, while the effectiveness of some lighting cues, such as relative brightness, were vastly inferior. Moreover, relative brightness, relative height, and relative size all significantly decreased in effectiveness with an increase in viewing distance.

Perception of virtual space in sparse environments has been shown to be poor, as predicted from geometrical considerations of the limited visual cues that are available.

Chapter 4 tested whether adding specific environmental cues (binocular cues, linear perspective, surface texture and scene clutter) could enhance perceptual performance in VR, with the aim of finding the degree to which individual cues contributed to a reduction in the amount of bias and variability in participant responses. It was concluded that adding differing visual cues did indeed have the hypothesised effect of improving performance, whereby a full-cue context allowed for the highest level of performance. This information is relevant to the current experiment as there will be sparse and cluttered environments tested: many visual cues will be observable in both conditions, however with the addition of numerous objects within the scene (clutter) also comes the added information from both the presence of the objects, and the extra shadows that they cast. As per the results of the previous experiment, where the inclusion of more visual features within a scene enhanced performance, the responses for an experiment where there is a plethora of additional visual information (a cluttered environment) should therefore also produce a high level of performance, measured by few errors. The current chapter builds on the results in the previous chapters by assessing specifically the contribution of cast shadows on the perception of relative distance.

Understanding the contributions of particular cues in specific applications is important in the design of effective image display systems. An important illustration of this in the use of augmented reality (AR) for use in surgical procedures. De Paolis and De Luca (2019) tested whether an AR 3D visualisation during a surgical procedure produced higher performance than when using simple 2D camera visualisations. Computed tomography scans were taken of the patient's area which was to be operated on; these were then meshed together and scaled to size so that a virtual representation of the organs and internal body structure were visible to the surgeon, through an AR device. A combination of visual and audio cues were then presented to the surgeon during the operation, with the hypothesis that this method would improve the intervention precision and avoid the risk of damaging anatomical structures. Presentation of depth was altered through a window visualisation (Bichlmeier & Navab, 2006), and visual text

information about the distance from the surgeon's tool, to the organ to be operated on. Audio signals were also provided as an additional distance cue to the specific organ. The results of these visual and audio distance cues provided high levels of accuracy and efficiency. In particular, the authors noted that shadowing indirectly contributed to depth perception, building on earlier work on the inclusion of shadowing as a cue to distance during surgery (Nicolaou et al., 2005). Despite its effectiveness as a distance cue, the authors nevertheless concluded that including shadows in their particular application would not be appropriate due to the complexity of the virtual scenes.

5.1.2 Shadows as a Visual Cue

Shadows occur when an object or surface disrupts the visibility of another surface or object to a light source, thus reducing the degree to which it is illuminated by that light source. The surface in shadow may belong to an object that is the same as or different from the one that is responsible for the shadowing. Mamassian et al. (1998) provide clear definitions of a number of related aspects of shading and shadows, including shading, attached shadows, and cast shadows. Shading refers to the variation in the amount of light reflected by surfaces, as a result of their orientation relative to the light source. This shading may be distinguished from the shadowing effect under consideration in this chapter, which result from the occluding of regions of surfaces from the light source. Attached shadows are those that are formed on the same surface that is occluding the light source. In contrast, cast shadows are those that are formed on a different surface to the one occluding the light source. It is cast shadows that are the focus of the current chapter, and in particular shadows that are cast on a horizontal surface (such as the ground plane or a table top) by objects that are not attached to the surface. This is shown in *Figure 5.1*. In this case, the observer is viewing an object at eye-height, and the object is illuminated from a single direction. When the light source is directly above the object, its shadow will be cast on the horizontal surface at the same distance as the object. If the light source if further away than that the object, the shadow will be

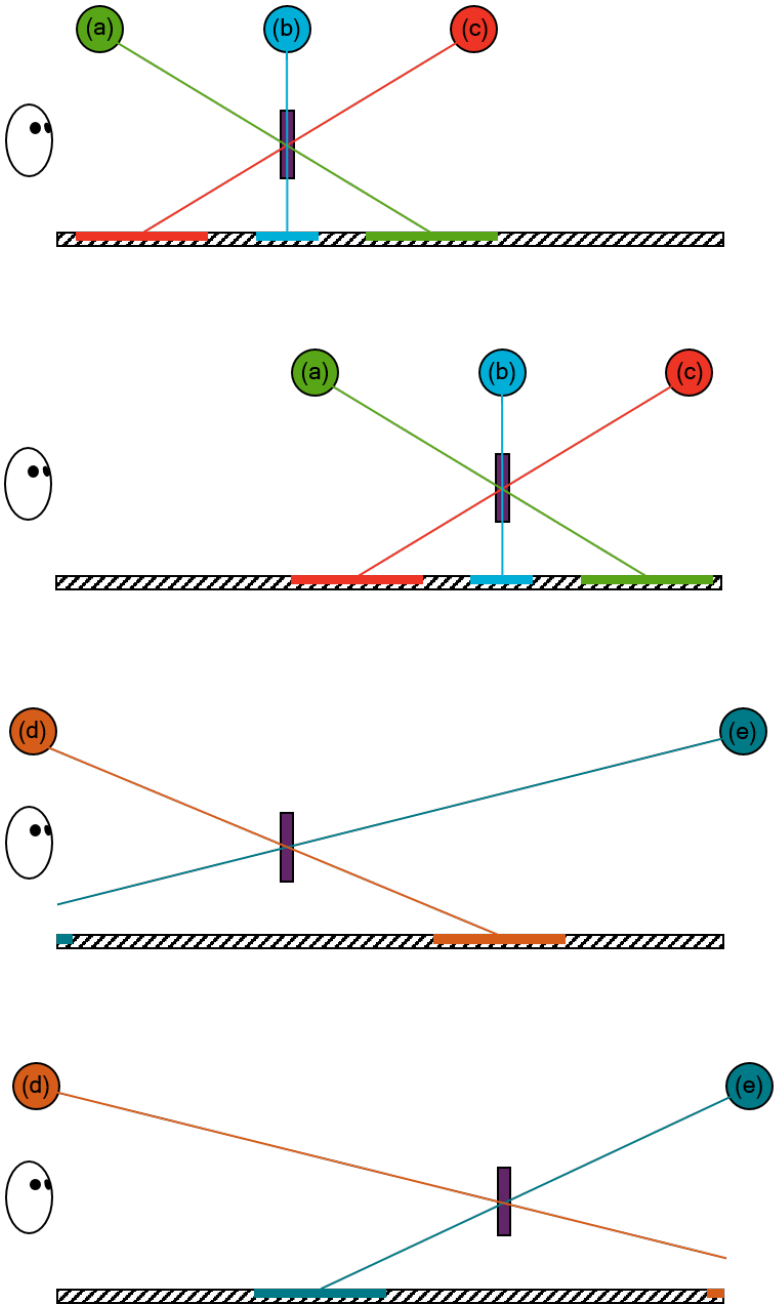


Figure 5.1: Examples of the cast shadows from the five different light sources used in this study, when the target object is in differing locations; shadow colour in accordance with light source colour. The different lighting conditions are: (a) `Locked in front` where the light is attached to the target, so shadows will keep a consistent shape and offset from the target; (b) `Locked on top` light again attached to target but there is no offset between the target and the cast shadow; (c) `Locked back` light locked to target in opposite direction of `Locked in front`; (d) `Static front` light source does not move in the environment, causing the cast shadows to vary in their size and shape when target location is adjusted; (e) `Static back` again static in environment but the opposite location of `Static front`.

cast at a distance that is nearer than the object itself. Conversely, if the light source is nearer than the object, or behind the observer, the shadow will be cast at a distance that is further away than the object. The location of the shadow on the surface provides information about the distance of the object relative to locations on the surface, albeit in a way that is ambiguous due to its dependence on the direction of the light source. Geometrically, it can be seen that, when the shadow is further away than the object, its distance from the object in the horizontal direction is given by:

$$\mathtt{D} = \frac{\mathtt{h}}{\mathtt{tan}\theta} \tag{5.1}$$

where D is the distance, h is the height of the object and θ if the direction of the light source. Thus, if the direction of the light source can be estimated, the distance of the object along the surface, relative to the location of its shadow, can be inferred.

It has been found that texture and cast shadows are used when determining the intensity and direction of a light source (Casati, 2004; Koenderink et al., 2004; Xia et al., 2014). Shadows are considered a pictorial cue, as they can be viewed monocularly and give an indication of depth even when presented on flat surfaces. Evidence suggests that pictorial cues, including casts shadow, can provide a more stable signal for depth and distance compared to binocular cues: Wanger et al. (1992) suggested shadows are a dominant pictorial cue for spatial and scaling tasks, whereas binocular disparity has been shown to decrease as distance increases (Allison et al., 2009; Palmisano et al., 2010). Wanger et al. (1992) used two tasks to assess the contribution of shadows to the perception of 3D surface layout. In the first task, they were presented with three objects lying above a ground plane. One object was resting on the plane and the other two were floating above it. The observer's task was to position the central object so that is was co-planar with the other two. It was shown that the presence of cast shadows, by helping to disambiguate the 3D location of the objects, improved performance on this alignment task. In a second, scaling task, observers were presented with two objects, one lying on

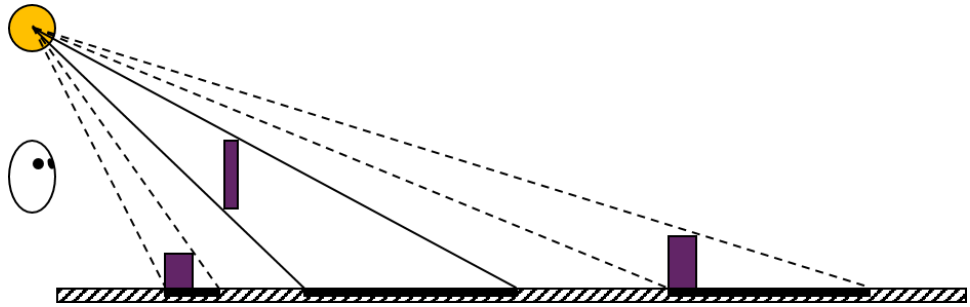


Figure 5.2: Examples of how the light source location can be more easily determined when there are multiple shadows from additional objects within the scene. Dashed lines identifying the light occluded by objects on the ground surface, where shadows and the objects are attached; solid line identifying the light occluded by the floating object, where the cast shadow is not attached.

the ground plane and the other floating above. Their task was to adjust the size of one of the objects so that the two matched in size. It was found that performance again improved when the floating object cast a shadow onto the ground plane that provided information about its three-dimensional location. Further evidence has also shown that the location of cast shadows influences the perceived depth and height of object, for which the effects can be observed in participants even as young as age three (Yonas et al., 1978).

Mamassian (2004) stated that because the correspondence of cast shadows to the objects that they belong to is obvious in simple, uncluttered environments, the shadow is therefore a reliable source of information for object location. To further this claim, te Pas et al. (2017) conducted an odd-one-out experiment, in which participants were presented with three stimuli with varying lighting and shadows. By collecting data of participant's eye fixations, reaction time, and percentage of correct answers the authors concluded that participants were mostly relying on the information given through the shadows and highlighting of objects, when making their coherence decisions.

Understanding properties of a stimulus is often much more accurate and precise when it is positioned centrally within the visual field of the observer. Nevertheless, Dee and Santos (2011) reviewed the literature on the visual effects of shadows and stated how shadows can provide information about objects even when they are outside the field of

view. They also stated that in simple cases on a flat surface, when a light source doesn't move, then physical properties of the object such as its size, motion, and shape can be inferred from the shadow, depicted in the simple example of *Figure 5.1*.

It has also been suggested that shadows can clarify ambiguous visual depth cues. Research by Kersten et al. (1997) showed that when presented with a ball stimulus moving linearly away from the observer in depth, and a shadow which would correspond with a bouncing ball, an observer was more likely to use the shadow information to perceive the ball as having non-linear movement. One of the conclusions of this study was that when a shadow moves, the observers are more likely to assume an object is moving, rather than the light source. In other words, the locations of shadows can be used to inform the estimation of the locations of objects under the assumption of a fixed light source.

Following on from this, there is evidence for the visual system completely ignoring certain visual cues in some circumstances. The expanding virtual room of Glennerster et al. (2006) provides evidence for observers failing to use the binocular disparity information provided throughout the study. Participants were situated in a room within VR and during the experiment were required to walk from one side of the room to another, along one edge. As they did this, the software continuously updated the environment in such a way that the room scaled up or down in size, in accordance with the participant's location within the room. Participants reported not noticing such a change in their surroundings, hence the conclusion that they were more willing to mis-estimate the distance they were walking, than actually acknowledge the changing binocular cues. This unconscious response to ignore binocular disparity was automatic in a very specific environment.

There are instances where specific lighting is used in combination with shadows to aid perception of a given environment. In cinematography, the use of lighting and shadows is commonly used to direct the observer to a point of interest, reflect the ambience within the environment, and even enhance the perception of the depths within a scene

(Adorama, 2021; MasterClass, 2021). Aside from the realm of entertainment, lighting cues are also used in flight simulations and the impact of different rendering techniques for the depiction of light has been explored (Nielsen, 2003; Ratzel, 2020), as well as for training doctors for laparoscopic surgery (McKenna et al., 2008; Shimotsu & Cao, 2007). The use of shadows is a widely utilised cue to aid performance in a number of active (training) and passive (entertainment) observer roles.

When an object is placed on a surface, its cast shadow will be directly adjacent to the object in the image, as shown in *Figure 5.2*. In this case, the size of the shadow, relative to the height of the object, can be used to provide information about the direction of the light source. For an object with a height (h), and a shadow with a length (f), the direction (θ) of the light source is given by:

$$\texttt{tan}\theta = \frac{\texttt{h}}{\texttt{f}} \tag{5.2}$$

5.1.3 Naturally Occurring Shadow Confusion in Pictures

In physical space there are occurrences where the shadows cast from objects are incoherent with the typical assumptions about the direction or shape of shadows. Lahaina Noon is the phenomenon where the sun passes directly overhead at the zenith at solar noon. This causes objects perpendicular to the ground to be void of any shadow in areas such as Hawaii (Hawaii News Now, 2020). Lack of shadows cause objects such as these to look either like they are poorly edited photos (when observing a representation of them), or give the (correct) perception that the light source is directly overhead (when observing in physical space).

Other naturally occurring distortions of perception from shadows come from the refractive properties of water: floating objects can often be observed to have shadows misaligned vertically, when the body of water is clear enough to see the floor. When the location of the sun is offset towards the left overhead for example, resulting shadows of

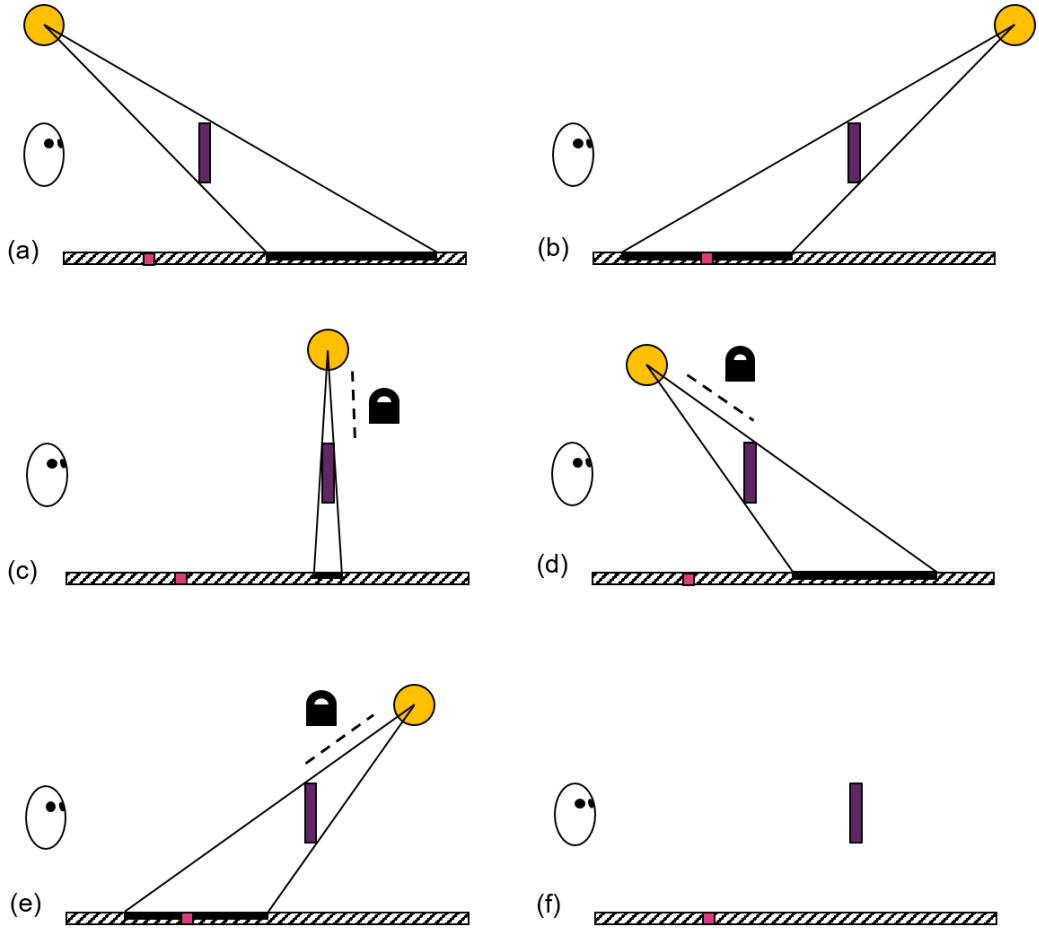


Figure 5.3: Examples of the shadow conditions where the light source is locked to the target (a) Static in front; (b) `Static back`; (c) `Locked on top`; (d) `Locked in front`; (e) `Locked back`; (f) `Ambient`. The light location moves in sync with the adjustments made to the target object location. Light source in three locations compared to observer, purple target object, pink reference line, black shadow on surface.

boats floating on the sea would be visible off-centre to the right. These offset shadows can be used to trace backwards and identify the location of the light source. The more boats which cast shadows with consistent offsets, the easier it is to identify the location of the light source casting these shadows. In this example, boats on the sea appear to float above the water because their shadows are not attached: this is due to the fact that the shadows are cast not on the surface of the water, but on the seabed, which is in fact at a distance below the surface. There is no visual information to specify the location of the surface of the water, at the same height as the boats.

5.1.4 Current Study

The experiment here consisted of two parts, both with a perceptual matching task presented to participants in a consumer VR headset. Participants were tasked with matching the distance of a floating cube to that of a horizontal reference line on a tabletop; this simple strategy is not typically available in real-world tasks. The light source was varied between blocks which, in turn, varied the cast shadow of the target. This paradigm was chosen because perceptual matching tasks have been shown to produce high-accuracy data (Rolland et al., 2002; Singh et al., 2010), and so the effects of both environment and lighting would be made clear.

In one condition, only ambient lighting was present, so that no shadows were cast. In all other conditions, one light source was tested, and its location relative to the object was varied. This light source was either in a fixed location in the environment so that it did not move with the object, or it was locked to the object, so that it moved and maintained a constant position relative to the object. In both cases, variations of the locations (and rotations) were used, with the light source either above (only when locked to the target), nearer (repeated for being static or locked), or further away (repeated for being static or locked) than the object and table top. These lighting conditions are depicted in *Figure 5.3*. The experiment was also repeated in two environments: a sparse and a cluttered environments, as detailed below.

The first hypothesis relates to the overall performance in the `Sparse` and `Cluttered environments`: (1) in the absence of shadows (`Ambient`) performance should be better in the `Cluttered environment`, due to the presence of additional visual cues to the three-dimensional layout of the scene.

In the `Sparse environment`, a number of hypotheses can be specified relating to the effects of the shadows. As shadows are expected to contribute to distance estimation, the first hypothesis here is that (2) performance will be worst overall in the `Ambient` condition. In addition, (3) performance is expected to be more precise for the conditions

in which the light source is directly above the target, when it is perfectly aligned with the reference, than when the light source is in front or behind. This is because in this condition, the observer is required to align two features that are overlapping, rather than when there is a spatial offset that will decrease the reliability of setting. It is also hypothesised that (4) biases will be smaller in this condition (`Locked on top`), since there may be a tendency to position the target object so that its shadow is aligned with the reference stimulus, rather than with an appropriate spatial offset given the direction of the light source. By having a direct comparison for the distance of the floating target, on a grounded surface such as the tabletop, the distances should be easier to estimate and therefore this condition can be considered as a baseline for performance. Furthermore, (5) performance will be more accurate when the light source is locked to the moving object than when it is static in the environment, since this creates a constant shape of the shadow, and a constant offset between the reference line and shadow when accurately aligned.

Finally, (6) it is hypothesised that any effects of cast shadows which improve the accuracy of settings will be more evident in the `Cluttered environment`, since this provides more reliable information about the direction of the light source, from the sizes of the shadows cast by the surrounding objects on the table-top.

5.2 Methods

5.2.1 Participants

There were 21 naive participants obtained from an online recruitment system as well as through word of mouth. Participants received a £5 voucher for approximately 45 minutes of their time to complete the study.



Figure 5.4: The `Sparse environment`, where the target object is floating above the tabletop, the reference line is visible on the tabletop, and the cast shadow is also visible.

5.2.2 Materials and Apparatus

An Oculus Rift headset and controllers were used, with two environments created in Unreal Engine 4.2. Shown in *Figure 5.4* is the virtual `Sparse environment` presented to participants: a room with a long table in and the target object. The architecture of the room, target, and reference line were created using the engine's Starter Content. Added into the `Cluttered environment` were the tabletop objects, shown in *Figure 5.5*. These were scans obtained from Patchs (2016), Dekogon Studios (2018), or scanned by the experimenter. A total of 21 objects were be positioned randomly on the 325 cm table surface.

The target object had a point light locked to the participant-facing side, with an intensity one lumen and attenuation radius of 10 cm. This light was programmed to not cast any shadows during the experiment.

5.2.3 Shadow Conditions

Both experiments contained the same six conditions, as follows and depicted in *Figure 5.3*:

- `Static in front` light source stationary, positioned at the edge of the table closest to the participant (position in environment: X = 37, Y = 0, Z = 175; rotation: X = 0, Y = -65, Z = 0), shadow size/shape changed as a function of the distance the target is from the light and is positioned between target and participant.

- `Static back:` light source locked to the edge of the table furthest from the participant (position in environment: X = 367, Y = 0, Z = 175; rotation: X = 360, Y = -65, Z = 180), shadow size/shape changed as a function of the distance the target is from the light and was positioned between target and far edge of table.

- `Locked on top:` light source attached to target object (position from target: X = 1800, Y = 0, Z = 580; rotation pointing directly down onto target), cast shadow was directly below target.

- `Locked in front:` light source attached to target object (relative transformation position from `Locked on top`: X = -600, Y = 0, Z = 300; rotation: X = 0, Y = -50, Z = 0), cast shadow moved with the target object and was positioned between participant and target

- `Locked back:` light source attached to target object (relative transformation position from `Locked on top`: X = 600, Y = 0, Z = 300; rotation: X = 0, Y = -50, Z = 180), cast shadow moved with the target object and was positioned between the target and the far edge of the table.

- `Ambient:` no point light source, no cast shadows.

The locked lighting conditions are shown in *Figure 5.3 (c), (d), (e)*; here, when the target object was moved by the participant, the light source would move accordingly as they were attached to one another. The light sources were sourced from Unreal Engine's Starter Content and were scaled at one along all axes.

The conditions were presented in blocks, each having 20 trial repetitions. In the



Figure 5.5: The `Cluttered` virtual environment, where additional objects are positioned randomly on the tabletop, casting their own shadows.

`Cluttered environment`, the objects on the virtual table cast their own shadows from the light sources; these were not be manipulated in any way.

5.2.4 Task and Procedure

The goal was for each participant to match the distance of the target object (a floating box) to that of a reference line on the table surface. Participants were told to read instructions and these were also described to them by the examiner; there was an opportunity to ask questions before starting. The order of experiments was counterbalanced, alternate participants were allocated to start in either the `Cluttered` or `Sparse environment`, however this was not disclosed to them. Their task was to align a target object with the reference line, irrespective of the condition they were in.

Once inside the VR headset participants were to sit in the indicated area and to stay there for the duration of the experiment. Once the experiment began, a horizontal line appeared on the surface of the table in front of them at random distances between 40 and 365 cm from the start location. The target object was positioned at eye-height and, using the controller's A and B face-buttons, participants could move the target along

the X axis (forwards/backwards in egocentric distance) to a position where it looked as though it was directly above the reference line. Once the distance was decided upon, the participant was to press the controller's trigger which saved the trial and moved onto the next one. The next trial automatically began by randomising the distance of the reference line and target box. No indication was given to participants about when the condition was to change.

The lighting direction (visualised through the shadow) was altered throughout the trials. The position participants set the target object to be in each trial was recorded, as was the accurate position of the reference line, so that the centimetre offset for each trial was able to be calculated.

5.3 Results

5.3.1 Formatting Data

The raw data output from the experiments were in a comma separated value format; each session had an independent .csv file named by the experiment name, time, and date of the session. Each row represented an individual trial with columns for the condition number, the distance of the reference line (in centimetres), and the set distance of the target (in centimetres). No reformatting of the data occurred, however an additional column was added to indicate participant number. A key was used to decipher which condition pertained to which condition number.

The centimetre signed and unsigned error for each trial was calculated by subtracting the distance of the reference stimulus from the set distance of the target stimulus. The unsigned error corresponded to positive, absolute values; while signed had both positive (indicating too far of a distance set) and negative (too short of a distance set) values.

The following exclusion criteria were decided upon before the experiment: any parti-

cipants who did not finish the entire experiment, and any participant who withdrew their data after a two-week window from completion. Additionally, removal of trials where the set distances were further or shorter than the length of the table was included, but this was not followed up (see *Section 5.4.3*). No data were excluded based on the exclusion criteria.

5.3.2 Analysis Strategy

A Pearson's correlation was conducted, separately for the two environments, on the distance of the reference line and set target positions in each trial to identify the general relationship between participant estimates and correct settings.

Signed error of estimates, where a positive value indicates a setting further from the observer than the reference stimuli and a negative value indicates a setting which falls short of the reference, were calculated by subtracting the distance setting from the reference distance in each trial. The amount of error is here used as a measure of the accuracy of estimates, and the sign of the error indicates where any systematic biases in results occurs. A t-test was then conducted on these values to compare overall performance between the environments.

Performance in terms of precision was quantified as the standard deviations of responses. Once again these were compared between lighting conditions through t-tests.

Linear mixed effect models were calculated so that the slope and intercept values could be retrieved from the data. Through the models used, slope values were able to indicate the amount of distance which participants were actually perceiving at each presentation distance of the stimuli; intercept values indicate the amount of offset, and therefore biases perceived. Here, performance in the lighting conditions was compared to that of the `Ambient` condition, as this was being used as a control condition as it was void of any shadows.

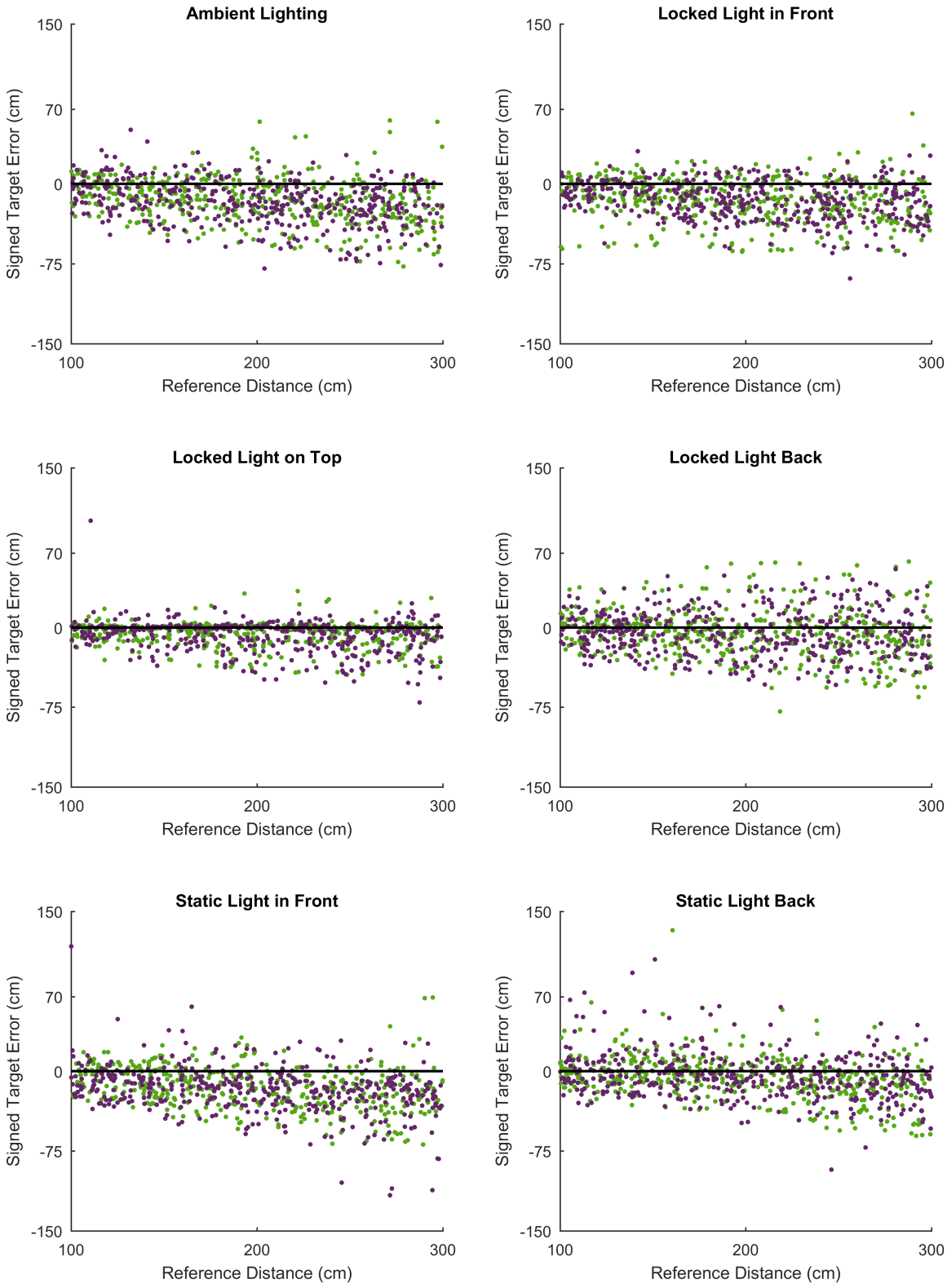


Figure 5.6: Scatter plot of the mean signed errors in the position of the set target in each condition and environment. Errors from the `Sparse environment` in green and `Cluttered environment` in purple.

5.3.3 Analyses

From the descriptive statistics shown in *Table 5.1* there was a consistent under-setting of distances, in that regardless of the lighting or sparsity of the context, participants set the target to be not as far away as it should have been. This trend can also be observed in *Figure 5.6*. A correlation was conducted on the distance of the reference line and set target positions displayed in *Figure 5.7*, to determine the overall relationship between the estimates and correct settings. In the `Sparse environment`, there was a strong positive relationship between the two (r = 0.935, p <0.001); the same trend found in the `Cluttered environment` (r = 0.933, p <0.001).

The effect of the environment was compared using a repeated measures t-test on the signed centimetre error in each trial. All data from all conditions were grouped together so that the effect of the environment only was observed. The mean error from the `Sparse environment` was -10.668 and that from the `Cluttered` was -11.284, which gave a non-significant difference overall between these environments (T (2519) = 1.441, p = 0.150, CL [-0.222 1.453]). Identifying the effect purely scene clutter had on distance perception, in the absence of shadows, was tested by comparing the accuracy of estimations in the `Ambient` lighting condition between the environments: again there was no difference between the extent of signed errors between the two environments (T (419) = 1.896, p = 0.059, CL[-0.076 4.213]). This suggests additional objects within a scene did not enhance nor diminish distance judgement accuracy.

The precision of responses, as measured by participant's standard deviation of responses in each condition, is presented in *Figure 5.8*. The variation in estimates across the conditions was examined: a repeated measures t-test was conducted on the standard deviation of the error in the `Ambient` condition between the two environments, to isolate and further compare the effect the scene clutter had on perception: T (20) = -0.132, CL [-2.449 2.157], p = 0.896. Having scene anchors alone did not have an impact on the variation of distance estimates.

Table 5.1: Descriptive statistics summary of the mean and standard deviation of the signed centimetre error in distance responses within each condition.

Condition		Mean	Standard deviation
Ambient	Sparse	-15.182	21.242
	Cluttered	-17.250	20.526
Locked in front	Sparse	-15.313	20.118
	Cluttered	-15.430	17.687
Locked on-top	Sparse	-6.258	14.055
	Cluttered	-7.762	15.007
Locked back	Sparse	-5.456	24.513
	Cluttered	-5.107	23.041
Static in front	Sparse	-15.205	19.669
	Cluttered	-17.473	22.269
Static back	Sparse	-6.595	20.746
	Cluttered	-4.679	22.044

A linear mixed effects model was used to analyse the data with the following equation, where the distance set, by participants (P) in each trial, is predicted by the distance of the reference line (D) and the lighting condition (L):

$$\texttt{DistanceSetting} \sim \texttt{D} + \texttt{L} + \texttt{D} * \texttt{L} + (1 + \texttt{D} \mid \texttt{P}) \tag{5.3}$$

The data for each environment was analysed separately: *Table 5.2* presents the slope and intercept information for the lighting conditions from both environments. Accurate performance would produce a slope value of one and an intercept of zero. In both environments the intercept was largest in the `Static back` lighting condition, and closest to zero in the `Locked in front` condition. All slopes were below one, indicating under-setting of distances. The closest slope estimate was produced by the `Locked on top` condition in the `Sparse environment`, and `Locked back` in the `Cluttered environment`.

As can be seen in *Table 5.1* and the unsigned error in *Figure 5.9*, the variation of estimates appears to be greater in some lighting conditions than others. The variation of responses, measured by the standard deviation of responses, revolved around 6 % of the distance tested; these data are presented in *Figure 5.8*. Averaging the standard deviation from both environments, it revealed that the environment with the largest

standard deviation was `Static back`. A repeated measures ANOVA on the standard deviation of the centimetre error, for each participant, revealed a significant difference in variation between conditions in the `Sparse environment`: ($F(5, 95) = 3.176$, $p = 0.011$), and for `Cluttered environment` ($F(5, 95) = 2.049$, $p = 0.079$). Planned t-tests on the two lighting conditions with the lowest standard deviations were undertaken, for each environment. The `Locked in front` and `Locked on top` were the two conditions with the smallest standard deviations, and it was found that in the `Sparse environment` the difference of precision between these conditions was not significant ($T(20) = 1.113$, $p = 0.279$, CL [-1.948 6.402]. In the `Cluttered environment` a non-significant difference was also found ($T(20) = 1.735$, $p = 0.098$, CL [-0.418 4.560]). The same was undertaken on the condition with the largest standard deviation, `Sparse Ambient`, and `Static back` for hypothesis (2): this was also found to be non-significant ($T(20) = 0.6846$, CL [-2.330 4.607], $p = 0.501$).

The centimetre error in distance settings for each trial can be represented by as a signed (directional) or unsigned (absolute) value. Signed errors indicate any biases in the direction of under- or over-setting, while the unsigned errors represents the magnitude of all errors, regardless of the direction. The same analysis as *Equation 5.3* was conducted once again, instead using the unsigned error as the dependent variable. The results for the two models on the unsigned error (the data from each environment analysed independently) are shown in *Table 5.3*, where accurate performance would be represented by intercept and slope values of zero. The unsigned error indicates which lighting condition or environment produced the most and least variable responses. Overall, the intercept estimates from the `Sparse environment` deviated the least from zero, and the `Locked on top` lighting condition also deviated the least.

The same analysis was conducted using signed error as the dependent variable, which is shown in *Table 5.4*. Here, the direction of any systematic errors are able to be detected: the slope estimates for both environments and all lighting conditions were negative values. Although the range from -0.041 to -0.115 is slight, the negative direction

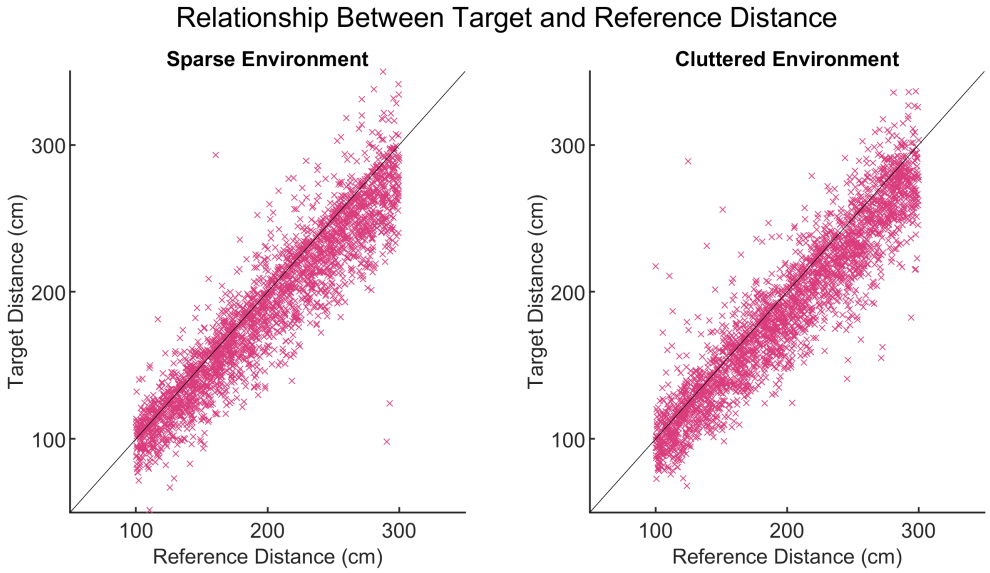


Figure 5.7: Scatter plot of the relationship between the distance of the reference line and the set target distance in each trial. Black line represents accurate performance with an intercept of zero and a slope of one.

reinforces the evidence of under-setting of distances, and therefore the perceived underestimation of presented distances in all conditions and environments.

5.4 Discussion

This experiment aimed to identify a lighting condition which produced the most accurate responses in a perceptual matching task, and whether the addition of scene anchors enhanced the estimates. A thin box was used as target so that the most visible face was flat and therefore did not have shading as an added visual cue. Clutter was added to the environment, however there was no change in accuracy of responses when presented with these added objects or in their absence. A light source behind the target, projecting a shadow positioned between the participant and the target, produced the most accurate estimates out of the lighting conditions tested.

5.4.1 Comparing to Hypotheses

The hypotheses were as follows: (1) `Cluttered Ambient` would allow for better performance than `Sparse Ambient`; (2) performance would be worst in `Ambient`; (3) precision

would be smallest in `Locked on top`; (4) biases would be smallest in `Locked on top`; (5) `Locked` conditions would produce more accurate estimates that the `Static` ones; (6) the lighting effects which improve the accuracy in the `Sparse environment` would be more evident in the `Cluttered environment`.

A t-test comparing the results of the `Ambient` condition of the two environments revealed no differences in the accuracy, nor precision: hypothesis (1) was therefore rejected. The addition of scene clutter did not enhance the perception of the layout of the scene, which can be seen by the similar distribution of results in both environments in *Figure 5.6*.

Hypotheses (2), (3), (4), and (5) related to the performance in the `Sparse environment`. The most inaccurate distance estimates were made in the `Ambient` and `Static in front` conditions for the `Sparse environment`. Neither of these lighting conditions were significantly worse than the other (*Table 5.3*), so (2) was able to be accepted, with `Ambient` being the joint worst lighting condition in terms of accuracy. Performance in terms of precision was measured by standard deviation, first a significant ANOVA found a difference within the conditions, then a t-test on the two with the highest standard deviations. Performance here was worst in `Locked back` as shown in *5.8*, however the difference between this and `Ambient` was not significant, also meaning (2) can be accepted for precision.

The condition which produced the smallest variability in estimates was the `Locked on top` condition, as predicted in (3). The t-test between the standard deviation of responses in this condition versus the second lowest standard deviation (`Locked in front`, *Figure 5.8*) revealed no significant difference.

Figure 5.9 presents the signed errors, the directional biases, for each lighting condition. From this it can be identified that in the `Sparse environment`, `Locked on top` and `Locked back` produced the smallest mis-estimation bias. This is further supported by the intercepts in *Table 5.4* whereby the intercept estimate for `Sparse Locked on`

`top` is the closest to zero (accurate performance). Due to this, (4) can also be accepted.

Hypothesis (5) was rejected, due to the inconsistencies in the average error obtained from the two `Static` conditions: `Static in front` resulted in a large amount of errors, however `Static back` resulted in much less: *Figure 5.9* presents this.

The slope estimates reported in *Table 5.2* present the level of mis-estimation of distance estimates in each environment and condition. Totalling up all slope values from the `Sparse environment` gives a mean average of 0.911, and 0.922 for the `Cluttered environment`. In the `Cluttered environment` the distance estimates were scaled at approximately 92 % of the correct values in this environment, which is 1 % better than in the `Sparse environment`. Further to this, analysing the difference in errors from the planned t-test of all results from the two environments was to determine whether (6) was to be accepted or not. This difference between the environments was not significant, therefore this hypothesis was rejected.

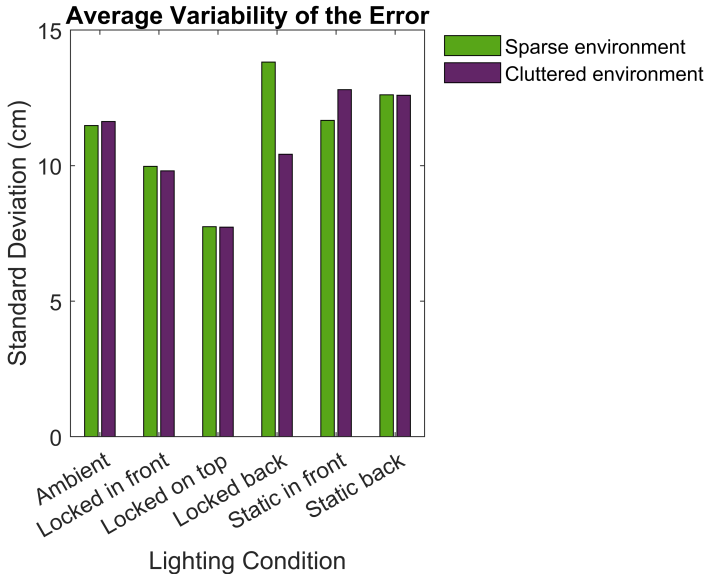


Figure 5.8: The average standard deviations on estimated from each condition, between participants. Responses from the `Sparse environment` in green and `Cluttered environment` in purple.

5.4.2 Interpretation of Findings

The mis-estimation of distances here is displayed by an error in setting the target stimulus to be at a shorter distance than the reference line. Clear instructions were given to participants and generally the trend shows that the task was completed correctly, so a fair assumption is that they understood the task was to align the target with the reference. Settings tended to be closer to the observer than they should have been in order to accurately align with the reference. This indicates an over-estimation of distance relative to the reference, if the closer target and farther reference were judged to be at the same distance. This means that perceptually, participants were perceiving the distance of the reference and the (shorter) distance of the target to be the same. They perceived the target to be further away than it actually was, so the space between themselves and this target was perceptually larger than in actuality, and also larger than the space between themselves and the reference. This explains the under-setting of distances found in this experiment.

This distance-matching task revealed errors in the distances being perceived in the virtual environment. Differences between the reference and target stimuli include the positioning along the Z axis, in terms of height in the visual field, and also realistic physics in the form of being grounded on a surface or floating. It would be wrong to conclude that all distances in the environment were being perceived in the same way, because the task itself required matching the distance of the target to the distance of a reference stimulus which was located on the tabletop. The errors in distance settings are evidence of some sort of distortion of distances which occurred in this scenario. The errors in the target's distance settings here have been identified as a psychological phenomenon where the participants mis-estimated the distance between them self and the distance they set. Alternatively, it might be that the reference stimulus was being perceived as being positioned at a shorter distance than it actually was, which resulted in this under-setting of the target: Ooi et al. (2001) determined that objects nearer a ground plane were perceived as being closer to observers than where they actually were located.

Instead of a distortion arising from the visual system, it could be argued that these findings are a result of the hardware itself. The headset used has hybrid Fresnel lenses which bend the light in a way in a way that is different to other stereoscopic set-ups. As a result of this curvature, aspects of the scene which are in the outer edges of the lens may be distorted differently from aspects in the centre of the lens. Although there are numerous possibilities for the cause of the mis-estimations of the target distance, the main conclusion from this study is that there was a relative difference in the perception of distances which was exacerbated and diminished by certain scene properties such as lighting and shadows.

The strong correlation coefficient between reference stimulus location and set target location suggests that participants were understanding the task and completing it appropriately. This coefficient does not however, indicate any systematic biases that may be present. Due to this, the mixed effects model was used to pinpoint the changes that the environment had on estimates, as well as lighting conditions. In all conditions, there was a consistent under-setting of the target stimulus, as can be identified in *Figure 5.9*. The regression outputs presented in *Table 5.2* show all positive slopes of the participant distance settings compared to the distances of the reference stimulus. This further corroborates the correlation output, that when the reference line increases in distance from the participant, the setting of the target stimulus also increases. However, accurate performance would produce an intercept of zero: the intercepts here show that in all conditions and both environments (except `Locked in front`) the results show a bias towards under-setting. The slope values show how well the participant estimates reflect the accurate performance, which would produce a slope of one. All slope estimates in both environments were below this, showing mis-estimation of the distances tested.

It is possible to compare the results found here to similar studies with different equipment and set-ups. One experiment by Swan et al. (2006) used a similar task in AR, whereby participants were to match the distance of a virtual target to a physical reference object, when observing the environment through a stereo optical see-through

display. The reference stimulus was positioned at different distances throughout the experiment, and the results showed errors in judgement increased with distance. The authors concluded that this linear relationship indicated that judgements were based on cues with linearly decreasing effectiveness. Much like the results from the current experiment, performance decreased with presentation distance. Other laboratory-based experiments using perceptual-matching tasks have found estimates of distance to be below veridical (Creem-Regehr & Kunz, 2010; Proffitt, 2006) and without judgement biases (Rolland et al., 2002). Comparing to other methods, Singh et al. (2010) revealed that open loop tasks such as blind reaching, resulted in less accurate judgements of distance than perceptual matching tasks.

No difference in accuracy nor precision was identified between the environments; however a difference was found between the lighting conditions, specifically between the conditions where the light was situated at opposing positions. The amount of overall error in responses for each condition can be seen in *Figure 5.6*, where the `Locked on top` lighting clearly produces the most precise distance estimates. This is supported by the average standard deviation of participants, depicted in *Figure 5.8* where the lowest standard deviation was found in this condition. There were the lowest amount of unsigned centimetre errors in `Locked on top`, but there were fewer signed errors in the `Locked back` condition.

Sugano et al. (2003) stated that shadows are an important depth cue in AR, but lighting needs to be consistent with physical lighting due to the see-through nature of the device. To add to this, a study by Gao et al. (2020) used an AR device to test the impact of shadows on distance perception. It was predicted that the distances would be under-estimated but also that virtual shadows would enhance accuracy in the perceptual matching task and that mis-alignment between physical and virtual lights would impact the estimates. The results confirmed that there was a mis-estimation of distances here and that lighting conditions also had a significant impact on these judgements, as well

as evidencing the impact of the mis-alignment. The results obtained between the two similar tasks support one another, despite differing head-mounted displays being used.

Observers typically assume that light comes from above and use this assumption to interpret how patterns of shading are related to 3D structure (Adams et al., 2004; Kerrigan & Adams, 2013). The results of this experiment suggest that when lighting is not positioned in this way there are more errors than when the light is shining from directly above. As can be seen from all graphs, there are large distortions in the judgements made in all lighting conditions and environments, however realistic lighting is important in virtual environments, if the applications or skills learned within the device are to be implemented into the physical world.

5.4.3 Evaluation of the Methodology

Initially, exclusion criteria were to include the removal of any trials with set distances outside the range of the table length, as this would not be possible should the experiment have been completed correctly. However this was not done, because in doing this, any extreme over-setting of distance when the reference line was positioned at the far end of the table, and under-setting when at the close end of the table that may have arisen from the effects of any shadows present, would have been eliminated.

Lighting conditions experienced in daily life tend to include ambient light from overcast weather diffusing the light source, and directional lighting from ceiling lights. The lighting conditions used here were created to replicate these, with overtly visible cast shadow effects (or not, in the `Ambient` lighting condition). Additionally, some less common light instances were created, which although not as common as static lighting in typical viewing, are used in specific implementations of VR which have been discussed in *Section 5.1.2*.

Surdick et al. (1997) stated that an object's relative brightness decreases with the square of the distance between the object and the point light source that illuminates that object. At close viewing distances, a given change in the simulated distance of the object would produce a relatively large change in its brightness. At farther viewing distances, however, that same amount of change in the simulated distance of the object would produce a much smaller change in the relative brightness of the object. The target object had its own light source, to counteract this phenomenon so that the lightness of the target was not an additional distance cue for the participants to utilise.

The reference stimulus was presented at random distances during each trial of the experiment. Having random distances between a minimum and maximum distance range is a more thorough methodological design than having categories of distances, such as near and far; or at a set number of distances with specific intervals. The random distances avoids the possibility for participants to get used to and remember the relative size of the target at each of the distances, and then try to replicate the size, rather than complete the trial as required.

The two environments were created so that a comparison of whether scene clutter aids in the perception of distances. The specific objects used in the `Cluttered environment` were chosen based on their commonality in everyday life, all varying in size and shape so that the shadows produced would also be varied. The positioning for these objects was random, a structured approach to the positioning of the clutter may influence the perception of the scene.

Anecdotally, it was noted that the speed of movement of the target object appeared faster in the `Sparse environment` compared to `Cluttered environment`. Two conclusions may be drawn from this, either that speed perception is affected by visual clutter, or that there was a lag in the software display caused by additional objects being rendered in the scene. The former would be interesting to test, as this would have applications for

the amount of information presented during situations where any lag would have dire consequences; to pilots training in virtual heads up display equipment, for example.

5.4.4 Future Research

AR combines virtual components to the observation of a physical environment, often through a head-mounted display similar to that of VR headsets. AR devices tend to be smaller and more portable than VR devices, as often they are not needed to be tethered to a computer. The findings of this study lead to a much related investigation on the impact of lighting and shadows when there are multiple light sources. Through the use of a beam splitter, virtual scene components can be positioned in such a way that they are viewed in conjunction to the physical environment. An investigation into the effects of mis-matched artificial lighting sources and physical light sources would be useful; because the use of AR devices is much more dynamic than VR and there is observation of physical space, the likelihood that virtual scene components are rendered exactly in accordance with physical lighting is lessened. The effects of virtual stimuli rendered with differing lighting conditions to the physical environment would be a useful and practical source of data for AR device manufacturers and/or software developers.

The application of these findings could be taken into consideration when creating puissant software. An example of this would be the advancements in software such as the security screening scanner by Linev Adani (2021); traditional body scans such as x-rays are used to detect foreign objects concealed within an individual, or internal injuries, whereby the virtual representation of the individual is presented on a screen for an examiner to inspect. Through the use of the artificial intelligence algorithm used in this technology, the guesswork is removed by adding virtual, artificial visual cues for the examiner to better identify abnormalities in scans. Although the background research behind the efficacy of this technology is unknown, it is assumed that the addition of these colour-coded cues enhances user detection rates, which supports the proposal here that the addition of visual cues in the current study (clutter, cast shadows) will lead to

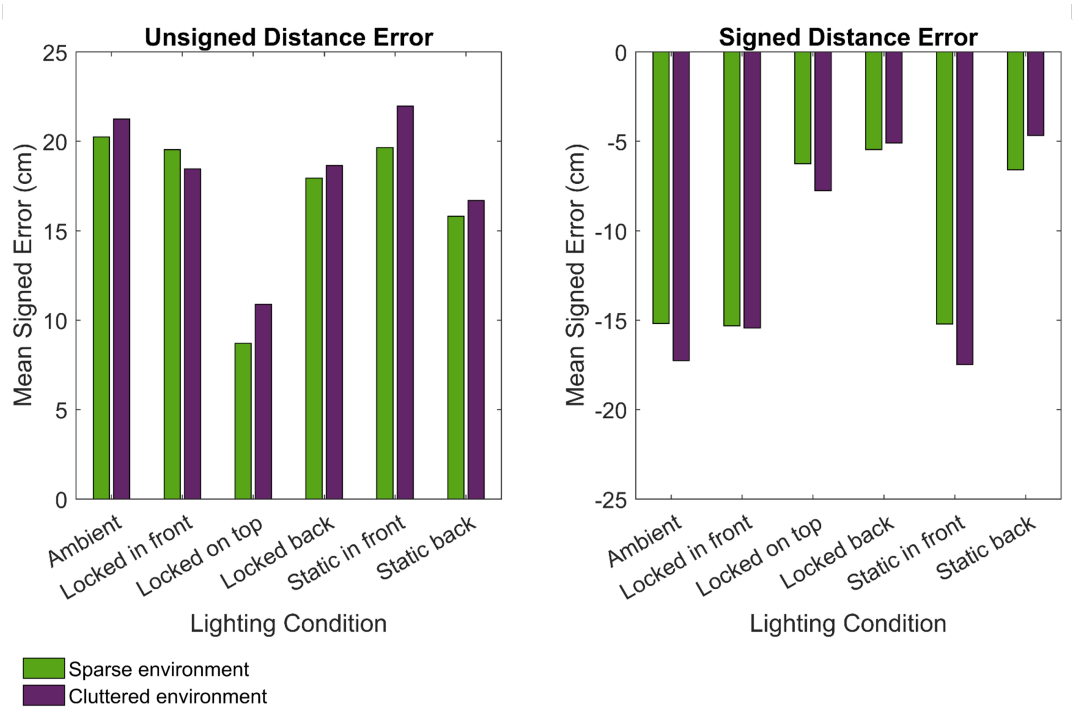


Figure 5.9: The mean error from the six lighting conditions. The average unsigned and signed error for the `Sparse environment` in green; `Cluttered environment` in purple.

higher accuracy rates.

Additional factors have been found to influence the perception of virtual environments: Slater et al. (1996) established that the greater the extent to which a participant can modify a virtual environment, the greater the resulting degree of immersion. The effect immersion has on perceptual learning is something which is yet to be tested: an experiment to investigate how immersion impacts accuracy of perceptual tasks would be able to be created using a similar paradigm as the current study. Here, immersion was not measured but including an immersion questionnaire (such as the Immersion Experience Questionnaire by Jennett et al. (2008)) as well as a condition where the participants have more control over the environment would allow for a directional relationship between immersion and perceptual accuracy to be observed. A greater extent of control over the environment could include movement of clutter objects on the table versus having no control over the locations of objects. It would be predicted here that with increased immersion, accuracy would also increase.

It would be beneficial to incorporate eye-tracking into a future version of the experiment to identify how often the shadow is observed compared to the target object. This, in combination with lighting conditions where binocular disparity is eliminated, would be beneficial for understanding how heavily the shadows cue is used during such a task.

Cluttered objects and an abundance of shadows on the tabletop may have been distracting for some participants, which may have contributed to the non-significant difference between the two environments. After discussions with some of the participants it became apparent that a subset of the participants valued the objects in the `Cluttered environment` differently than other participants. Some of the individuals suffered from attention deficit hyperactivity disorder and claimed that the additional objects in the `Cluttered environment` were too distracting, and found the task more difficult in this environment. A portion of the participants who did not have the disorder reported the cluttered objects to be useful in their distance judgements. From these informal discussions, it would be interesting to repeat this experiment with two groups of participants: neurotypical and neurodivergent, to explore the mechanisms of spatial awareness and the way this differs between the those with an attention deficit disorder and those who do not. It would be predicted, from these discussions, that the neurodivergent group would have increased performance in the `Sparse environment` compared to the `Cluttered environment`, whereas the neurotypical would show the reverse trend.

5.4.5 Conclusions

Adding clutter to an environment, where there are more shadows cast and more points of reference, does not enhance the accuracy of distance estimates within the scene compared to an identical environment minus the clutter. Clutter objects are therefore not used as points of reference, or there are subsets of participants who use these as anchors while others found them to prevent adequate completion of the task.

In accordance with previous literature, there was a consistent under-setting of dis-

tances throughout the experiment. Lighting conditions where the cast shadow of a target were positioned between the target and the observer reduced the under-estimation of distances perceived in virtual environments.

From the lighting techniques tested, overall there were the fewest errors in the `Locked on top` lighting condition and most in the `Static in front`. With regard to the variability of responses, the `Locked on top` condition also produced the most precise estimates, and `Static back` produced the least precise.

Some of the lighting conditions presented in this experiment can be considered to be 'unnatural' in typical viewing of physical space, as there are few instances where a light source is attached to an object with some sort of offset. The two static lighting conditions are more representative of typical viewing. However, implementations of unnatural lighting can be seen in specific applications of VR and cinematography. It is in these instances where the perception of distance can be enhanced through the use of the conditions presented in this experiment.

Table 5.2: Regression Analyses of the distance perceived in each lighting condition in the two environments, using Equation 5.3. Accurate performance would have an intercept of zero and a slope of one. The p value indicates whether the row is significantly different from the values of the `Ambient` condition.

Condition		**Intercept Estimate**	**SE**	**T**	**Intercept CLs**	***p***
Ambient	Sparse	6.199	3.698	1.676	[-1.052 13.449]	
	Cluttered	5.822	4.112	1.416	[-2.241 13.855]	
Locked in front	Sparse	-0.485	3.162	-2.114	[-6.685 5.715]	0.035
	Cluttered	-0.986	3.387	-2.010	[-7.629 5.656]	0.045
Locked on top	Sparse	2.987	3.115	-1.031	[-3.121 9.095]	0.303
	Cluttered	3.865	3.364	-0.586	[-2.732 10.452]	0.561
Locked back	Sparse	10.129	3.132	-1.031	[3.988 16.270]	0.210
	Cluttered	3.120	3.298	-0.819	[-3.347 9.587]	0.413
Static in front	Sparse	7.238	3.123	0.333	[1.114 13.362]	0.739
	Cluttered	5.536	3.308	-0.086	[-0.950 12.022]	0.931
Static back	Sparse	16.005	3.106	3.157	[9.914 22.096]	0.002
	Cluttered	9.749	3.271	1.201	[3.3349 16.163]	0.230
		Slope Estimate	**SE**	**T**	**Slope CLs**	***p***
Ambient	Sparse	0.891	0.020	44.311	[0.852 0.931]	
	Cluttered	0.885	0.017	52.008	[0.852 0.919]	
Locked in front	Sparse	0.925	0.0152	2.250	[0.895 0.955]	0.025
	Cluttered	0.930	0.016	2.783	[0.898 0.961]	0.007
Locked on top	Sparse	0.952	0.015	4.048	[0.923 0.982]	<0.001
	Cluttered	0.943	0.016	3.551	[0.911 0.974]	<0.001
Locked back	Sparse	0.922	0.015	2.044	[0.892 0.951]	0.001
	Cluttered	0.959	0.016	4.616	[0.928 0.990]	<0.001
Static in front	Sparse	0.887	0.015	-0.266	[0.857 0.917]	0.026
	Cluttered	0.885	0.016	-0.045	[0.854 0.916]	0.964
Static back	Sparse	0.888	0.015	-0.222	[0.858 0.917]	0.824
	Cluttered	0.928	0.016	2.704	[0.897 0.959]	0.007

Table 5.3: Regression Analyses of the unsigned error in distance settings for each lighting condition in the two environments. Accurate performance would have an intercept and a slope of zero. The p value indicates whether the row is significantly different from the values of the `Ambient` condition.

Condition		**Intercept Estimate**	**SE**	**T**	**Intercept CLs**	***p***
Ambient	Sparse	-2.152	2.790	-0.771	[-7.624 3.320]	
	Cluttered	4.679	2.779	1.684	[-0.769 10.128]	
Locked in front	Sparse	4.599	2.668	2.531	[-0.633 9.830]	0.011
	Cluttered	4.186	2.853	-0.173	[-1.409 9.781]	0.863
Locked on top	Sparse	1.572	2.628	1.417	[-3.581 6.726]	0.157
	Cluttered	-0.404	2.834	-1.794	[-5.960 5.153]	0.073
Locked back	Sparse	1.395	2.642	1.342	[-3.787 6.576]	0.180
	Cluttered	10.277	2.778	2.015	[4.830 15.724]	0.044
Static in front	Sparse	-0.797	2.635	0.514	[-5.964 4.370]	0.607
	Cluttered	5.056	2.786	0.135	[-0.407 10.520]	0.892
Static back	Sparse	5.067	2.621	2.754	[-0.072 10.206]	0.006
	Cluttered	13.414	2.755	3.170	[8.012 18.817]	0.002
		Slope Estimate	**SE**	**T**	**Slope CLs**	***p***
Ambient	Sparse	0.112	0.016	7.082	[0.081 0.143]	
	Cluttered	0.083	0.015	5.633	[0.054 0.112]	
Locked in front	Sparse	0.075	0.013	-2.877	[0.050 0.101]	0.004
	Cluttered	0.070	0.014	-0.995	[0.043 0.097]	0.320
Locked on top	Sparse	0.037	0.013	-5.913	[0.012 0.062]	<0.001
	Cluttered	0.055	0.014	-2.066	[0.029 0.082]	0.039
Locked back	Sparse	0.083	0.013	-2.311	[0.058 0.108]	0.021
	Cluttered	0.042	0.013	-3.070	[0.016 0.069]	0.002
Static in front	Sparse	0.103	0.013	-0.704	[0.078 0.128]	0.481
	Cluttered	0.084	0.013	0.055	[0.058 0.110]	0.956
Static back	Sparse	0.053	0.013	-4.670	[0.029 0.078]	<0.001
	Cluttered	0.016	0.013	-5.080	[-0.010 0.042]	<0.001

Table 5.4: Regression Analyses of the signed error in distance settings for each lighting condition in the two environments. Accurate performance would have an intercept and a slope of zero. The p value indicates whether the row is significantly different from the values of the `Ambient` condition.

Condition		**Intercept Estimate**	**SE**	**T**	**Intercept CLs**	***p***
Ambient	Sparse	6.199	3.678	1.676	[-1.052 13.449]	
	Cluttered	5.822	4.112	1.416	[-2.241 13.885]	
Locked in front	Sparse	-0.485	3.162	-2.114	[-6.686 5.716]	0.035
	Cluttered	-0.986	3.387	-2.010	[-7.629 5.656]	0.045
Locked on top	Sparse	2.987	3.115	-1.031	[-3.121 9.095]	0.303
	Cluttered	3.8651	3.364	-0.582	[-2.732 10.462]	0.561
Locked back	Sparse	10.129	3.132	1.255	[3.988 16.270]	0.210
	Cluttered	3.120	3.298	-0.819	[-3.347 9.587]	0.413
Static in front	Sparse	7.238	3.123	0.333	[1.114 13.362]	0.739
	Cluttered	5.536	3.308	-0.086	[-0.950 12.022]	0.931
Static back	Sparse	16.005	3.106	3.157	[9.914 22.096]	0.002
	Cluttered	9.749	3.271	1.201	[3.335 16.163]	0.230
		Slope Estimate	**SE**	**T**	**Slope CLs**	***p***
Ambient	Sparse	-0.109	0.020	-5.417	[-0.148 -0.070]	
	Cluttered	-0.115	0.017	-6.730	[-0.148 -0.081]	
Locked in front	Sparse	-0.075	0.015	2.250	[-0.105 -0.045]	0.025
	Cluttered	-0.071	0.016	2.723	[-0.102 -0.039]	0.007
Locked on top	Sparse	-0.048	0.015	4.048	[-0.077 -0.018]	<0.001
	Cluttered	-0.057	0.016	3.551	[-0.089 -0.026]	<0.001
Locked back	Sparse	-0.078	0.015	2.044	[-0.108 -0.049]	0.041
	Cluttered	-0.041	0.016	4.616	[-0.072 -0.010]	<0.001
Static in front	Sparse	-0.113	0.015	-0.266	[-0.143 -0.083]	0.790
	Cluttered	-0.115	0.016	-0.045	[-0.146 -0.084]	0.964
Static back	Sparse	-0.112	0.015	-0.222	[-0.142 -0.083]	0.824
	Cluttered	-0.072	0.016	2.704	[-0.103 -0.041]	0.007

Chapter 6

Exploration of Perceptions Surrounding Virtual Reality Technology

User experience has been suggested to be a crucial predictor of future use of goods or services. For this reason, the current chapter collected qualitative and quantitative data which delved into both experienced users and novices about their perceptions of virtual reality equipment, the concerns surrounding data collected by the equipment, and facets needed to be taken into consideration for future developments. The data was collected via an online questionnaire, which was publicised on numerous social media platforms and online noticeboards. It was revealed that data security is still an issue which has yet to be overcome. Individuals are aware of the risks associated with the newer technology, in terms of privacy and physiological side effects. The opinions surrounding precautions and important VR features from the public do not differ drastically from the opinions of manufacturers, however this study gives an in-depth perspective into the improvements desired by the public.

6.1 Introduction

6.1.1 Virtual Presence and Immersion

Virtual reality (VR) provides many benefits in comparison with other display technologies through its ability to create immersive experiences for users. A key goal of VR is the creation of *presence*, defined as the phenomenological sense of being in an environment (Gibson, 1979). This sense of being in the virtual environment is central to the definition of presence, and includes not only the sense of spatial presence, which is the experience of being physically present in an environment, but also the dimensions of social presence, realism, and engagement (Berkman & Akan, 2019). Social presence is the feeling of being present with a real person (Oh et al., 2018), while realism refers to the extent to which the virtual experience corresponds to the experience of an equivalent physical environment (Lombard & Jones, 2015; Slater, 2003), and engagement requires a strong experiential connection with the elements of the virtual environment (Lombard & Jones, 2015).

Presence is often discussed alongside the closely linked, but less consistently defined, concept of immersion (Berkman & Akan, 2019; Slater, 2003). Immersion focuses on the technical components of the virtual reality hardware and software, and the degree to which they are able to create the perceptual experience required for a sense of presence. Immersion also addresses the extent to which the user is enveloped in the virtual world, and the corresponding isolation from the real, physical environment (Witmer & Singer, 1998).

When defining and evaluating a VR experience, Steuer (1992) emphasises the importance of focussing on the user experience, rather than the technology used to create that experience. In so doing, he emphasises the central role played by *vividness* and *interactivity* in creating an immersive experience. Overall, the properties of perceptual immersion in a virtual environment, and the reactivity of this perceptual experience to the users movement and behaviour, are critical in creating the sense of presence that sets the experience of VR apart from the experience of viewing other types of displays. Con-

sequently, the ability of designers to freely create immersive, responsive environments has huge potential application in many fields.

While science fiction often features VR as a futuristic technology, capable of transporting the user to new locations, in practice it has already been adopted in order to achieve this goal in a number of industries, not limited exclusively to the gaming or entertainment sectors. The purpose of the current study was to identify current levels of experience with, and attitudes towards virtual, augmented, and mixed reality (XR umbrella term for these types of devices), in order to identify possible barriers to its widespread adoption, and the desired future direction of its potential users.

6.1.2 Benefits and Applications of Virtual Reality

Being *immersed* in an environment or activity is described by Douglas and Hargadon (2000) as being completely absorbed within the ebb and flow of a familiar narrative schema. Brown and Cairns (2004) suggested that, although interconnected, presence is an important factor in creating the goal of strong immersion. This is a different use of the concept of immersion from that which specifically characterises VR technology, as outlined above, and is an example of the many and varied ways that overlapping terminology is used, for example in the creative arts (Berkman & Akan, 2019; Bjørner et al., 2016; Jarvis, 2019; van Dam et al., 2020). In this context it is primarily the ability of VR to create presence that adds to the absorption in the experience. While immersion and presence are often evaluated using questionnaire measures, Slater (2004) and Sanchez-Vives and Slater (2005) argue that this is more meaningfully defined and measured as the extent to which participants respond realistically to virtual events and situations. While VR aims to encourage users to feel both present and immersed in the virtual environment, a sensation of having *ownership* of virtual avatars and *agency* over controlling these avatars are also important factors contributing to one's ability to become fully engaged in another world (Mohler et al., 2008; Yu et al., 2019).

VR has huge benefits over other display technologies in creating immersive experi-

ences: manufacturers and novice users revel in this immersion that the headsets provide, which is why the equipment has been successfully used in a multitude of applications in preference over other displays. There remain a number of technical considerations which continue to limit the degree of immersion that can be achieved in VR (Hibbard, 2022 in press; Scarfe & Glennerster, 2015, 2019). A limited field of view (Howlett, 1990; Prothero et al., 1995), tracking lag (Held & Durlach, 1989; Tharp et al., 1992), limitations in the resolution, dynamic range, and rendering accuracy of the graphics (Sugano et al., 2003), and limited interaction with the environment (Slater & Steed, 2000; Slater et al., 1995) have all been found to impact on the ability to become immersed in the virtual environment, or to limit the sense of presence. In the context of the visual experience within VR, many of these cues influence the perception of the 3D layout of the environment, including the perception of distance (Thompson et al., 2004; Willemsen et al., 2004) and depth (Livingston et al., 2009; Zhang & Hua, 2010). As these properties are important contributing factors to successful immersion and presence within virtual environments, VR devices strive to achieve perceptual realism and accuracy. Wide fields of view providing a substantial binocular overlap, and an enclosed headset ensures that the user cannot engage visually with the physical world. Close to real-time tracking of the user's position means that the displayed information can be updated to reflect their movement around, and interaction with, the environment. These and other technological improvements are all ways that manufacturers of headsets have tried to improve the perceptual experience of VR.

Given the improvements in the quality and usability in VR in recent years, the potential applications of the technology are vast. To date, VR has been targeted primarily at the gaming market through advertising (Game, 2016; Oculus, 2017). While this may not be explicit, it is telling that both of these example trailers from leading manufacturers exclusively contain content from gaming and entertainment experiences, and that in a product launch presentation virtual reality experiences are referred to exclusively as *games*. This might create the impression that the predominant market for VR is gaming. In fact the ambition for this technology is that, like mobile phones, it will become

ubiquitous in all our personal and professional social interactions, as summarised by the Chief Scientist of Facebook Reality Labs (Abrash, 2014):

> We're working on technology that will change not just computer gaming, but potentially how all of us interact with computers, information, and each other every day. I think it's going to be the biggest game-changer I've ever seen – and I can't wait to see how far we can take it.

Indeed, the applications of VR already extend beyond gaming. Within the sphere of entertainment, for example, VR has been successfully used in theatre to augment the traditional live theatre experience, allowing people to enjoy a version of that experience from their own homes (Jarvis, 2019; van Dam et al., 2020).

VR has also been used extensively for workplace training. For example, it has been used in skills training, such as developing motor skills for surgeons training for surgery (Seymour et al., 2002), and in training firefighters (Narciso et al., 2020) and other emergency services workers, allowing them to develop their skills in a safe environment (Bodyswaps, 2021; FLAIM, 2021).

Entertainment, education, and training are well-known applications of immersive environments; distraction from pain and other symptoms, and rehabilitation of hospital patients are examples of a less well-known application. With the lifetime risk of being diagnosed with any form of cancer at approximately 39.5 % (Howlader et al., 2019), there is a definite need for research into interventions that can reduce the associated negative effects. A meta-analysis by Chirico et al. (2016) examined nineteen VR-based experiments on cancer patients to distinguish whether or not VR was an efficient intervention to relieve some of the psychological symptoms and improve emotional well-being during treatment, if there is a specific type of experience which has the best efficacy, and the relevant features of VR in this case. This systematic review, abiding by the relevant guidelines (Moher et al., 2009), identified an initial 803 potentially relevant

articles through key word searches, which was narrowed down to 19 studies that were read for analysis. A downfall here is that studies may have been removed which potentially contained important information pertaining to the aims of the analysis. In addition, the different types of methodology used between experiments poses questions for the reliability of results: some studies used self-report techniques while others relied on more empirical forms of data collection so comparison of these results may have been problematic. The overall findings of the paper, however, were that VR was an effective mechanism through which to relieve stress of cancer patients during treatment, with the best results found when used in combination with chemotherapy as opposed to before or after treatment.

VR has been used in the rehabilitation of other ailments in addition to cancer. Jack et al. (2001) used a VR setup, with feedback gloves, to rehabilitate hand function in stroke patients. With costs lowering for higher performance personal computers and VR equipment, the authors suggest the future of out-patient care may be more successful with the use of programmes similar to the one they tested. Looking at movement of the hand in terms of range, speed, fractionation (ability to move one finger independent of the others) and strength, the rehabilitation programme included three to four sessions each day for nine days total. Although there were only three participants, so results can not necessarily be generalised, the data as a whole showed improvement in these four domains from the first trial to the last. Supporting this, a meta-analysis by Laver et al. (2012) evaluated the effects of VR and interactive video gaming compared with results from either control groups or alternative rehabilitation techniques for stroke patients. Out of the 19 studies looked at, only seven showed significant improvement in functioning. On the surface this does not seem impressive, however further analysis revealed extremely low rates of the traditional adverse effects associated with VR (motion sickness, headache), suggesting that using VR is relatively safe as an alternative to traditional methods of stroke rehabilitation, but the outcome may depend on external factors such as the software or task itself. It is important in future research to identify the attributes of the VR experiences and programmes associated with successful out-

comes.

In order for this broad range of applications of VR to be realised, it is important to understand how this technology is perceived not only by current consumers, but potential future users. This allows us to appreciate both the positive drivers, and potential barriers, to the acceptance and embracing of the potential of VR.

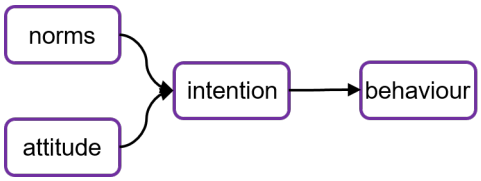


Figure 6.1: The Theory of Reasoned Action (Fishbein & Ajzen, 1977).

6.1.3 Acceptance Models for Behaviour

Many theories have been created in the hope of explaining human behaviour. The Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1977) models the relationship between an individual's attitude toward a certain behaviour, and perceived pressure toward performing the behaviour, to their behavioural intention. Separation of the behavioural intention from actually carrying out the behaviour is important here, as it was revealed that merely having the intent to do something does not mean an individual will do it. Nevertheless, it was found that one's intention to engage in a specified behaviour was the best predictor of whether or not they would engage in that behaviour. These intentions are influenced by the societal norms and their own attitude, as illustrated in *Figure 6.1*. A meta-analysis performed by King and He (2006) revealed these factors to be the best predictors of behaviour, although further research into the theory have since found other factors to be also be important for adopting different types of behaviour. Importantly, factors such as the accessibility to the environment which is necessary to complete the behaviour are not included; no matter how much an individual wants to partake in a behaviour such as playing a new game, if they do not have access to the equipment that allows the behaviour to be completed, then the behaviour is categorically impossible.

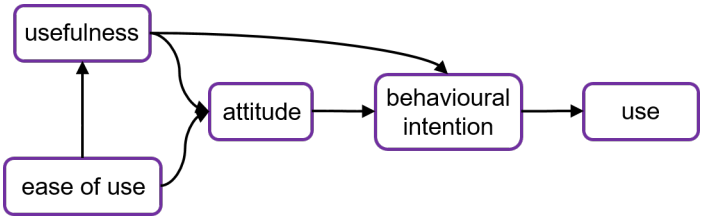


Figure 6.2: The Technology Acceptance Model (Davis, 1989).

This research is discussed as follows.

The Technology Acceptance Model (TAM) (Davis, 1989) was created specifically as a development of the TRA that has been used to understand how people accept and use technology. The perceived usefulness and ease of use of a technology were found to both be significant indicators of system use (*Figure 6.2*). Here, *perceived usefulness* was defined as how much the individual believed that using the piece of technology would enhance their job performance, and *perceived ease of use* was the amount the individual thought the technology would be effortless. Consistent with this, the meta-analysis by King and He (2006) also found perceived usefulness to be the strongest predictor of intention to use. Interestingly, attitudes and societal pressure were not included in this model, and the predictors of acceptance here are more directly linked to the behavioural usage of the equipment, rather than indirect societal or other pressures the individual may face.

TAM has been validated in banking and commerce (Bruner II & Kumar, 2005; Chau & Lai, 2003; Ha & Stoel, 2009; Hernandez et al., 2009) and has been extended to fit a range of different technologies (Amoako-Gyampah, 2007; Choi & Ji, 2015). Stelter et al. (2020) stated that acceptance models such as the TAM lack important variables, or are not fully applicable to many circumstances. The extent that an individual is actually able to engage in the behaviour due to resources and opportunities is not taken into consideration in either the TAM or the TRA. In particular, when it comes to using technology that inherently comes with significant financial costs, the perceived ability to adjust behaviour to accommodate the equipment would also be a sensible inclusion for

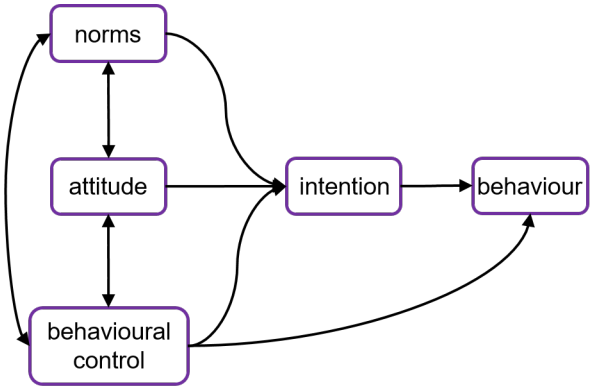


Figure 6.3: The Theory of Planned Behaviour (Ajzen, 1991).

the model. The TAM has been found to only account for around 40 % of a system use (Legris et al., 2003), however since 2003 there have been extensive advancements in technology, so the actual percentage may be very different with the adoption of technologies such as VR. Additionally, humans engage in behaviour which might not always be considered as useful; the entertainment sector encompasses activities and experiences which can be deemed as enjoyable, rather than useful. Intentions are complex and should be re-specified as being dependent on the end goal and why it is being requested. For example, personal intentions (wanting to try a new game so can join in conversations with friends) or collective intentions (incorporating a programme in the workplace), which are both linked to interpersonal influence. Thus, the Theory of Planned Behaviour (Ajzen, 1991) was modelled as a more comprehensive theory than the TAM as it incorporated this as a predicting factor.

The Theory of Planned Behaviour (TPB) is shown in *Figure 6.3*. On the surface, this is an extension of TRA with many of the same predictors such as attitude, norms, and intention. Critically however, perceived control over one's behaviour is added here. The absence of this feature can be considered a limitation of the TRA, as being able to literally partake in an action. This was therefore added to the model with the goal of better explaining behaviour.

6.1.4 Acceptance Models for Virtual Reality

The ambition of VR companies is to integrate headsets into daily life, so questions arise from marketers and developers regarding the long-term return on investment, including how easy the consumers find handling the equipment; whether this predicts actual usage; and the public's likelihood of purchasing the equipment. The Virtual Reality Hardware Acceptance Model (VR-HAM, *Figure 6.4*) by Manis and Choi (2019) is an updated version of the TAM, which aimed to objectively assess the TAM specifically for VR, while adding extra concrete variables of age and price, as well as subjective ones including curiosity, usefulness, enjoyment, and ease of use. Contrary to the TAM, perceived enjoyment was found to be the most powerful predictor of the acceptance of VR, consistent with research on handheld devices (Bruner II & Kumar, 2005). This suggests there may be a greater connection between an individual's beliefs and their use of mobile devices and VR headsets, than VR devices and other types of technology where acceptance models have been applied, such as augmented reality (AR) (Elshafey et al., 2020; Leue et al., 2014) and mechanical aids (Graham et al., 2009; Moyon et al., 2019).

Once again however, interpersonal influence is not included in this model. The influence of an authority figure pressuring a change in behaviour in the workplace will likely impact use of a device.

In comparison, Lee et al. (2019) surveyed 350 participants when analysing a variation of the TAM to understand the acceptance of VR. Perceived enjoyment, social interactions, and the strength of social ties were added to the model. It was found that social interactions and the strength of the social ties increased perceived enjoyment. Further to this, perceived enjoyment had a more significant effect on intention to use than perceived usefulness. This is an important point of difference, since perceived usefulness is the main element of the TAM.

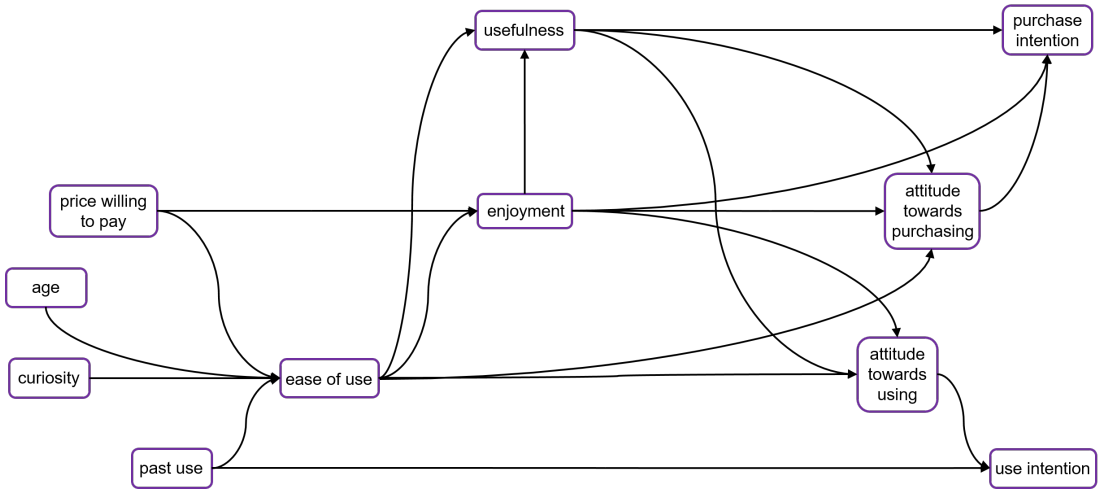


Figure 6.4: The Virtual Reality Hardware Acceptance Model (Manis & Choi, 2019).

Approaching this concept from a different perspective, Bertrand and Bouchard (2008) tested the original TAM, with perceived cost as an added predictor, on 141 participants who were already interested in using VR. With this participant sample, it was found that the intention to use was predicted only by perceived usefulness. This suggests that there are variables such as one's background knowledge, hobbies, and interests that are key factors when aiming to predict future behaviour, and that the important factors for the acceptance and adoption of VR will differ across different groups.

A study by Sagnier et al. (2020) noted that many of the acceptance models do not include subjective factors such as emotions, aesthetic appeal, and hedonic qualities; this is important as there has since been evidence showing the impact that these can have on intent to use technology (Davis et al., 1992; Van der Heijden, 2004; Venkatesh & Speier, 1999). It is therefore necessary to gather more qualitative information from consumers on their opinions before attempting to predict behaviour.

6.1.5 Current Research Questions

The goal of this study was to understand current attitudes towards VR and related technologies, with a view to how these relate to its adoption, particularly in applications beyond gaming. To do this, a novel questionnaire was developed, which included both

forced-choice and free-text answer questions. Commonalities between individuals who did and did not at the time of answering choose to use VR were pinpointed, and their explanations as to why were also assessed. This method was used to understand whether the various acceptance models relating to VR, such as TAM and VR-HAM, take the most important factors into consideration, now that the technology is more widely available and more people may have some experience with its use compared to when these models were created.

There was predicted to be a difference in the factors deemed important by those who are current users of the technology, compared to those who did not use it. No demographic differences were expected between users and non-users.

Links between the degree of enjoyment and engagement with other types of gaming and VR usage were expected, as well as a negative relationship between VR usage with propensity for motion-sickness and the presence of existing visual or auditory impairments. Perceived usability was also expected to have an impact on the enjoyment of VR.

In addition to questions on the usefulness and user experience of the technology, potential concerns were addressed, pertaining to user-privacy. There were probing questions on the viewpoints of the types of data that can be collected by devices, which may raise themes on potential concerns over user monitoring, data privacy and behaviour, and mood modification through the use of personalised content (Kumar & Sharma, 2014; Verma, 2014). All of these factors are important considerations for the individual, and societal, acceptance of XR technology. If individuals are not confident in the privacy of their own data, if non-users have concerns about covert or overt video and audio capture, or if there is a sufficient lack of trust in the way that companies select and personalise content, then the technology is unlikely to be accepted.

6.1.6 Mixed Methods Design

Traditionally in psychology research, quantitative data are collected through empirical studies,and statistical analyses are performed on these numerical data. Where it is possible to frame research questions so that they can be answered using quantitative methods, this approach is important in identifying robust trends.

On the other hand, qualitative data is often used in political and social science research, in studies where the reasoning behind behaviours or activities are being looked at, through the analysis of speech or written text. Qualitative information can be collected via open-ended questions, in written format such as questionnaires, or interviews and focus groups. Additionally, observational studies can also provide this type of data.

Quantitative data can provide the *what* of a participant's answer, whereas qualitative can be more enlightening in delving into the theoretical question of *why*, for example, particular attitudes or intentions are held. Multi-methodology designs can be used to collect both data types in one hit. In the current study, there was exploration of opinions (through open-ended text questions), structured rankings (through ordinal scales), and background demographics (dichotomous closed-questions) of the participants taking part. By collecting both types of data, a more comprehensive discussion of the results was aimed for, than what could have been given if only one data type was used.

Sentiment analysis was used to analyse the data for some of the free-text answers. Sentiment analysis is a way to detect opinions within text, and group these opinions together based on similar meanings. For example, 'happy' and 'delicious' would both have a positive sentiment, whereas 'dirty' and 'foul' would have a negative sentiment. The sentiment lexicon is the most important resource for most analysis algorithms, as this is where words and their sentiments are stored (Feldman, 2013). Sentiment analysis is often used when analysing data surrounding politics (Agarwal et al., 2018; Kušen & Strembeck, 2018), current affairs (Dahal et al., 2019; Manguri et al., 2020), and sports (Gong et al., 2021; Wunderlich & Memmert, 2020). Data to be analysed in this way

needs to be in written format, but this can of course include transcriptions from verbal accounts.

6.2 Methods

6.2.1 Setup and Procedure

The questionnaire was created in Qualtrics, with five blocks of categorised questions. A total of 50 possible questions were included, which are presented in *Appendix A*.

Participants were first presented with the study's information and consent form, two forced choice options for consenting appeared: if participants did not consent the study was programmed to end with a thank you note. Those who did, were able to access the questionnaire.

For the ease of structuring, the questionnaire was sectioned into blocks. During the analysis, the questions were categorised by theme: questions pertaining to the user's *background* (demographics and information to create a profile), *experiences* (history with the technology), *security concerns* (privacy of data), and *opinions* (exploring viewpoints through free-text responses).

6.2.2 Data Collection

Opportunity sampling was used by posting the questionnaire on the University of Essex's Department of Psychology online participant recruitment forum. Students were rewarded 0.5 course credits for taking part. Additionally, the questionnaire was publicised on Twitter, LinkedIn, and through word of mouth in order to recruit a broader range of participants. Participants were also encouraged to spread the link to others who may have been interested in the topic.

The overall number of data sets collected was 178, with 158 arising from the University's online participant recruitment system and 20 from external sources. However, 34 were excluded from the analyses as not representing full and unique sets of responses.

There were 14 who started but did not complete the questionnaire, and 20 represented cases where the participant completed the survey multiple times. In these latter cases, only the first attempt was included in the analysis.

6.2.3 Analysis Strategy

All data analysis was performed using R 4.0.3 and RStudio 2021.09.0. Percentages and frequencies for each possible response were calculated for the responses to the closed questions, as appropriate.

The McNemar's Chi Squared test (RDocumentation, 2021a) was conducted on the binary response questions, as these questions produced dichotomous responses from the participants. This allowed testing for whether the responses were evenly distributed across the two categories. The script was as follows:

$$\begin{gathered}\texttt{TEST_DATA} < -\texttt{DATASET\%} > \%\texttt{select(FACTOR1, FACTOR2)} \\ \texttt{TEST_TABLE} < -\texttt{table(TEST_DATA\$FACTOR1, TEST_DATA\$FACTOR2)} \\ \texttt{dimnames(TEST_TABLE)} < -\texttt{list(FACTOR1} = \texttt{c(“No”, “Yes”), FACTOR2} = \texttt{c(“No”, “Yes”))} \\ \texttt{mcnemar.test(TEST_TABLE)}\end{gathered} \tag{6.1}$$

Quantitative data analysis was conducted on the answers to the Likert questions, using the Wilcoxon signed rank test with continuity correction, `wilcox.test` from the package RDocumentation (2021b). This test was used due to the 5-point ordinal nature of the Likert questions. Here, the one-sample test was conducted against the middle value of 3 (labelled as being 'neither disagree nor agree' with the question statement). The test therefore indicates cases in which the median responses differ from this neutral, central category. The equation is as follows:

$$\texttt{wilcox.test(FACTOR, mu} = 3, \texttt{alternative} = \text{“}\texttt{two.sided}\text{”}) \tag{6.2}$$

For each of the free-text questions, one of two qualitative analyses techniques was

conducted. As part of the pre-processing, filler words were removed using:

$$\texttt{as_tibble(map(TERM.MATRIX, \sim discard(.x, \sim .x\%in\%stopwords(“en”))))}$$

as were punctuation and numbers. The removal of words such as 'an', 'I', 'the' allowed for the qualitative analysis techniques to identify semantically meaningful words instead of these commonly used terms. Different types of qualitative analysis was then undertaken on these passages.

Sentiment analysis uses natural language processing and text analysis techniques to identify subjective information from text (Hussein, 2018). The analysis can be used on a document as a whole (Yessenalina et al., 2010), a sentence (Farra et al., 2010), or words individually (Engonopoulos et al., 2011). The polarity of sentiment (positive or negative) behind the words or phrases is detected through this method of analysis; here, the analysis was conducted on the word level. Sentiment analysis was implemented on questions 8, 17, 24, and 26; using the `comparison.cloud` from the `wordcloud` package (Fellows, 2018). The frequency of individual words was first calculated, then the positive and negative associations for each word were retrieved from the package. The words were sorted into their sentiment category, and the frequency of each word directly impacted the size of that word in the wordcloud: higher frequency resulted in a larger font size. The words needed to have a frequency of more than two to appear in the cloud. For the standard wordclouds presented, only the frequency of individual words within and between participant responses were counted and plotted, using the same package.

Additionally, correlated topic model analysis was used from the `topicmodels` package (Grün & Hornik, 2011). In this form of analysis, topic proportions were identified in the free-text answers and correlations were identified after implementing a logistic normal distribution (Aitchison, 1982). The algorithm then derived a variational inference (Blei & Lafferty, 2006). The words associated with each of the found topics are displayed as a graph, with the beta weightings identifying the probability of a word occurring in a given topic (term distribution). The number of themes set for the algorithm to detect

was intuitively decided upon, as was the number of terms within the theme to display. Higher beta weightings suggest a word is more likely to appear in the text. This analysis was conducted on question 28.

6.3 Results and Discussion

6.3.1 Background Profile

A total of 144 individuals completed the questionnaire. Demographic information shows the gender split was 88 female, 55 male, 1 other; age ranged from 18 to 61 (mean = 23.44, standard deviation = 7.50); 69 had 20/20 vision or better. Of those who did not have 20/20 vision, 53 wore glasses, 22 wore contact lenses, and 15 did not know the accuracy of their vision as they had never had an eye test.

Generations are commonly categorised by date of birth, although there are some conflicting dates depending on the source. Generally speaking, the *Silent Generation* are those who lived through World War 2; *Baby Boomers* were those born during the surge of births after the war; *Generation X* were born when technology was quickly advancing, but was not readily available; *Millennials* are the digital natives as they were born when the world migrated from analogue to digital; *Generation Z* have never known a life without technology due to its mass availability (Colby & Ortman, 2014). The results here show that the most common generation identified with was Generation Z (n = 61, followed by Millennial with n = 41).

Only 10 of the 144 participants had ever had a test to determine whether they were able to perceive 3D stimuli. From the participant pool, the expected number of participants who are stereo-blind was 11 (7.6%) (Chopin et al., 2019) which would have an impact on their perceptual experiences of VR technology, and consequently also the responses in this questionnaire.

Additional information, such as ailments and impairments, were explored: 46 parti-

cipants (32 %) experienced travel-induced motion sickness and 22 (15 %) experienced technology-induced motion sickness. Visual impairments (not including wearing glasses or contact lenses) were experienced by 11 (8 %), and 7 (5 %) experienced auditory impairments. A McNemar test (*Equation 6.1*) was conducted to determine whether individuals who experienced visual impairments enjoyed using VR. Out of the eight individuals who had tried VR and experienced some sort of visual impairment, 87.5 % enjoyed using VR despite their impairment. The test suggests that there was a statistically significant difference between those who did and did not enjoy VR, varying with visual impairment (N = 77.108, df = 1, p <0.001).

The majority of the participants reported enjoying the theatre (n = 131, 91 %), as well as watching television (n = 141, 98 %), and playing board games (n = 120, 83 %). Other activities which were enjoyed by the majority were online (n = 98, 68 %) and offline gaming (n = 106, 74 %). The most popular console used to play games was reported to be mobile phone/tablets: due to the portability of these devices and how phones are usually kept on one's person throughout the day this is no surprise. The least common was a Mac computer; *Figure 6.5* shows how people are more likely to not play virtual games at all than to play on a Mac. The impact of these results pertain to the potential applications of VR which may or may not take off in the future. Information on the activities that people who are interested in the technology enjoy can be used to guide future developments, for instance watching television was enjoyed by majority of participants, which may be why Netflix have expanded into the realm of VR (Netflix, Inc., 2018).

6.3.2 Virtual Reality Experiences

Only a minority of participants had not tried VR before (32.63 %). Of these, the most likely reason for not having tried VR was the lack of opportunity (*Figure 6.6*). Other, less common themes were not being able to afford it, and not having any interest in try-

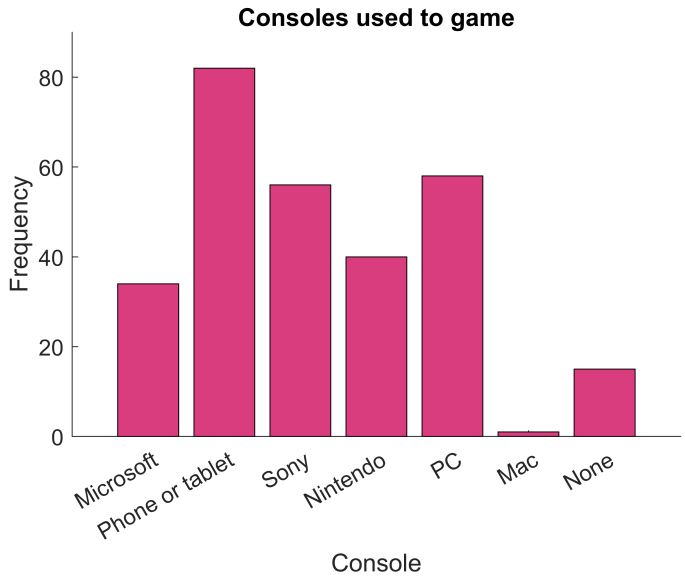


Figure 6.5: Frequency of the device brands used for gaming. With *Microsoft* referring to the Xbox series; *Sony* the PlayStation series; *Nintendo* the Wii, Switch, and DS variations.

ing the technology. Additionally, one participant gave the reasoning that the technology *"doesn't appeal to me. Don't like the idea of not knowing what's going on around me whilst having a headset on"*. This suggests there was a minority of individuals who prefer less immersive devices, which is contradictory to the goals of VR headsets. Those who did not enjoy the immersive experiences that VR has to offer are unlikely to purchase or use the technology, so any developments made on their behalf are moot.

AR and VR differ in the intended purposes of the headsets: VR presents an entire virtual environment to the observer, while AR adds virtual components to the otherwise normal viewing of the real world. Due to this difference, AR tends to have more day-today, functional purposes that can be integrated into other tasks, while VR is currently more geared towards entertainment and learning, where an immersive experience is required. Mixed reality (MR) also another, similar type of technology, where virtual components of a scene are added to the viewing pf a real environment, much like AR. When asked about their usage of other types of devices, 65.27 % said they had not tried AR and 88.19 % said they had not used MR. The total who had used all three (VR,



Figure 6.6: Word cloud for the free-text responses to the question 'For what reason(s) have you not tried VR before?'. Larger words indicate higher frequency.

AR, MR) amounted to only 9.02 % of the participants. The most common segue device appeared to be VR, to then lead on to trying the other devices, only 6.25 % had not tried VR but had tried either AR or MR.

Surprisingly, the most common generation to use AR were the Millennials, with 13.19 % of those who had tried AR being from this generation. This is surprising because of mobile applications such as Snapchat being common among Generation Z, but this generation category ranked second with 12.5 %. An explanation for this could be that Millennials have had more time to try it out compared to Generations Z. The younger generations have experienced the uprising of this technology at a key point in their lives, compared to the older generations: the Silent Generation for example would have less of a vested interest in using this technology, as the younger generations value technology as being more useful than the older ones (Venkatesh et al., 2003). Additionally, older generations have been found to experience more technology-related anxiety than the younger ones (Chung et al., 2010; Czaja et al., 2006; Ellis & Allaire, 1999) which may also act as a deterrent from using it.

Of the 97 who had tried VR, only 4 reported using it often, 14 used it sometimes, and 79 used it rarely. There were no prompts in this question so the participants were

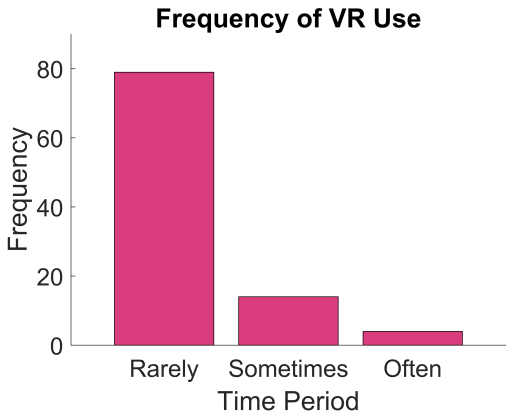


Figure 6.7: Frequency of how often participants use VR.

to decide for themselves what these categories of time meant and can be seen in *Figure 6.7*. This appears at odds with the findings of Manis and Choi (2019), who reported that 50 % of their participants used VR weekly. A related question was also asked to those who had reported to trying VR before, which was when the last time participants used VR (*Figure 6.8*): only one person reported to have used VR the same day as taking the questionnaire, and only one person had used it the day prior. For the other response options on when they had last used the technology: a few days ago n = 6, a week ago n = 4, a few weeks ago n = 9, over a month ago n = 33, and for 'so long ago I can't remember when it was' n = 43.

Of those who had tried VR, 91.75 % enjoyed using it. The most common of the pre-set uses for VR was for recreational purposes (66.10 %), followed by education (16.10 %), and the least common reported option was for business purposes (2.54 %) which are displayed in *Figure 6.9*. Interestingly, 15.25 % of individuals responded in a free-text answer for other purposes, however these included 'entertainment', 'gaming', and 'fun' which could be argued to come under the theme of recreation. A use of research appeared seven times, which again could be included under the pre-set theme of business.

A mere 11 out of the 97 who had tried VR before adjusted their VR device to their personal distance between their eyes or pupils, yet 24 reported to having had this distance measured previously. It is important that this is recognised, despite the oppor-

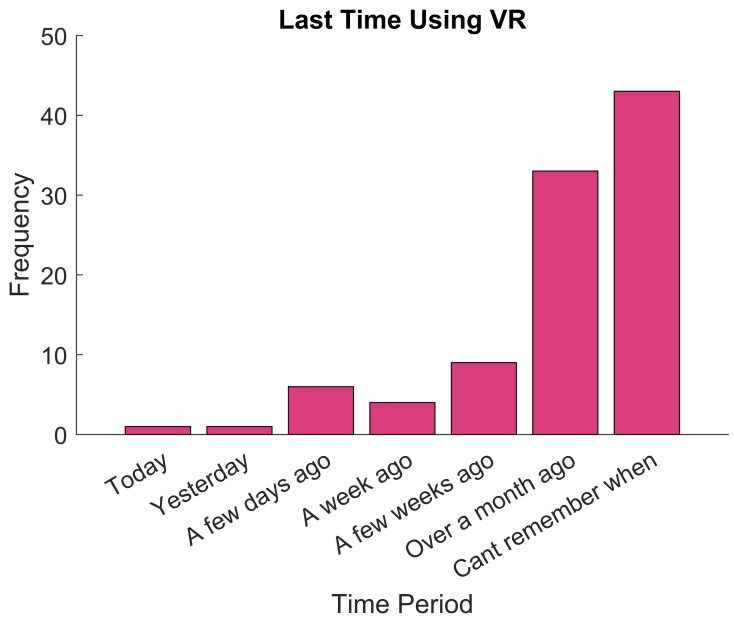


Figure 6.8: Frequency of the last time participants used VR.

tunity to set the screens to match the observer's distance between their eyes (IPD) in many headsets, due to the problems in the realism and the comfort of the VR experience that can accompany an incorrect IPD (Hibbard, 2022 in press; Hibbard et al., 2020). Moreover, of those who reported to needing to wear glasses, only 14.70 % wore them while using VR (23.52 % reported sometimes wearing them).

Only 14.58 % of the sample owned their own VR device. The Sony PlayStation VR was the most popular (n = 6); followed by the Oculus Rift (n = 4); joint third with n = 3 was the Oculus Quest, Samsumg VR, HTC Vive; Oculus Quest 2 and Rift S with n = 2; HTC Vive Pro with n = 1. Only three people owned more than one device, no one owned three or more.

The available content was voted as the most important factor for VR, closely followed by the purchase price. *Figure 6.10* depicts this, along with the less popular responses of how customisable the technology is, and its popularity. Popularity could be argued to be related to the social norms predictor in the TRA and TAM. Both these models found it to significantly predict usage, in contrast with the low importance attached

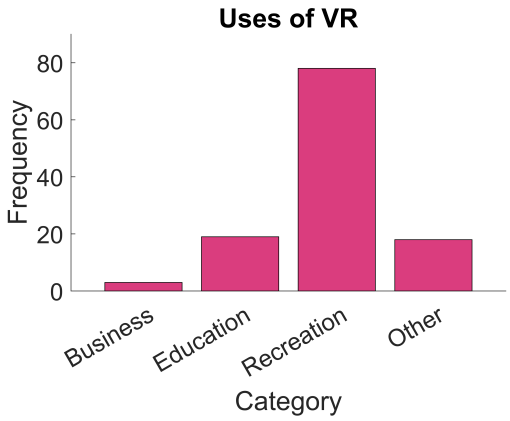


Figure 6.9: Frequency of the reported uses of VR equipment.

to it in responses found here. This could be due to the societal changes since these two models were created. A study by Hamamura (2012) suggested that in recent years society is becoming more individualistic, potentially as a consequence of economic growth (Bernstein, 1971). If this is the case, as a society there is no reason the growth will not continue and so the social norms and popularity of technology may be deemed an even less important aspect for new technology.

Figure 6.11 plots the most and least common responses to the Likert questions which asked participants whether or not they agreed with eight statements. There were no extreme modal values ('strongly disagree' nor 'strongly agree') for these questions. The results of the Wilcoxon signed rank tests (*Equation 6.2*) are shown in *Table 6.1*. These results show that VR device were perceived to be easy and enjoyable to use, necessary, but also perceived to be bulky. This is somewhat at odds with Simon et al. (2020), who found perceived usefulness as being the most important factor. Although rated significantly more positive than neutral, when referring to *Figure 6.11* it is apparent that participants viewed VR headsets as being only somewhat useful.

Perceived ease of use has been found to have a direct effect on intention to use (Agarwal & Prasad, 1997), however other studies found this effect to actually be mediated by perceived usefulness (Davis et al., 1989b). Here, both ease of use and usefulness were not considered to be any of the most important characteristics of VR, as shown in *Figure 6.11*.

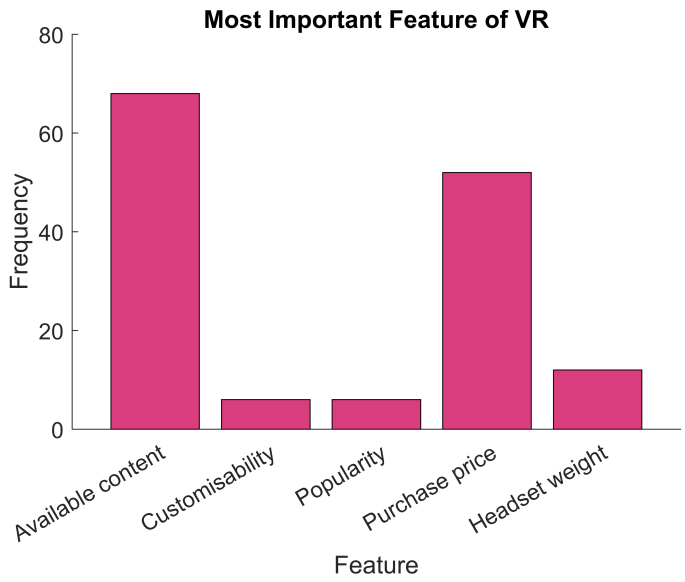


Figure 6.10: Frequency of the option chosen to be the most important feature for VR.

When asked to rank the importance of four different sensory features of VR devices (*Figure 6.12*), having accurate head-tracking was rated the most important, closely followed by visual graphics. The features that participants ranked the least important for VR devices were high quality sound and haptic feedback from controllers.

A meta-analysis by Van der Meijden and Schijven (2009) found that when using VR to train surgeons, the results did show positive findings which were in favour of force feedback, although many were ambivalent and not unanimous. It was suggested that haptic feedback is important during the early phase of psycho-motor skill acquisition. Because most of the users in the current study reported their uses as being more in the realm of entertainment (*Figure 6.9*) this may explain why haptics was ranked so low on the importance question; perhaps as psycho-motor skills are not being taught or learnt by these users, it is not as important.

When asked to rank the options of: the available content in VR, customisability, popularity, purchase price, and weight of the headset in terms of most to least important, the highest ranked feature were the content available, followed by the price.

A study by Jang et al. (2021) used 292 in-service teachers to explore their willingness

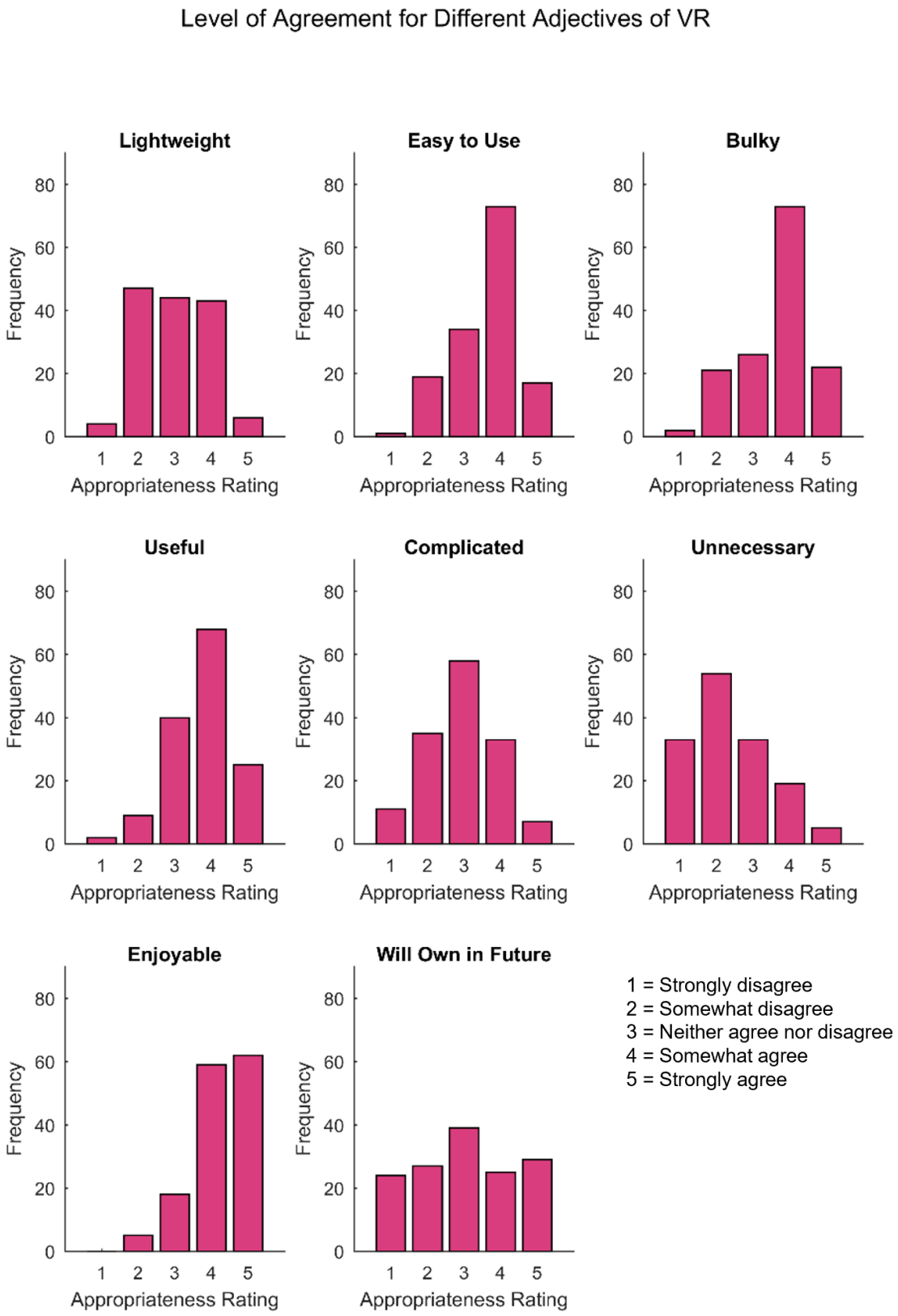


Figure 6.11: Frequency of the ranked appropriateness of different descriptions of VR devices.

Table 6.1: Output from the one-sample Wilcoxon signed rank test on the 5-point likert style question 'what are your thoughts on the VR headsets themselves?'.

Prompt	V	p	Direction
Lightweight	2529.5	0.988	Neutral
Easy to use	5120	<0.001	Positive
Bulky	5810.5	<0.001	Positive
Useful	4927	<0.001	Positive
Complicated to set up	1681	0.385	Neutral
Unnecessary	1165.5	<0.001	Negative
Enjoyable to use	7838.5	<0.001	Positive
I will own one in the next few years	2953.5	0.574	Neutral

to integrate AR and VR technologies in their teaching practices: technological pedagogical and content knowledge were found to have a significant influence over perceived usefulness and perceived ease of use. However, motivational support had a more direct impact as this affected perceived ease of use, which then influenced their attitudes towards technology use and behavioural intention to use. In this example, motivational support would be related to the support of the teaching establishment, however this could be related to popularity amongst peers. In this instance, popularity was ranked even less important than the weight of the headset, which can be seen in *Figure 6.10*.

6.3.3 Concerns Associated with the Technology

Participants were asked to explain any precautions they believed needed to be taken when using VR. The free-text responses were analysed using topical analysis. The analysis did not produce any strongly correlated terms for the themes, however the four retrieved topics generally make sense for the question that was asked. The terms and themes are displayed in *Figure 6.13*. The themes that the algorithm detected were interpreted as being: (1) physiological risks; (2) privacy; (3) the disconnect between virtual and physical space; (4) awareness of a secure environment.

Theme 1 was predicted to appear, and does appear in many of the responses throughout the rest of the questionnaire. Sickness and damage are negative associations with

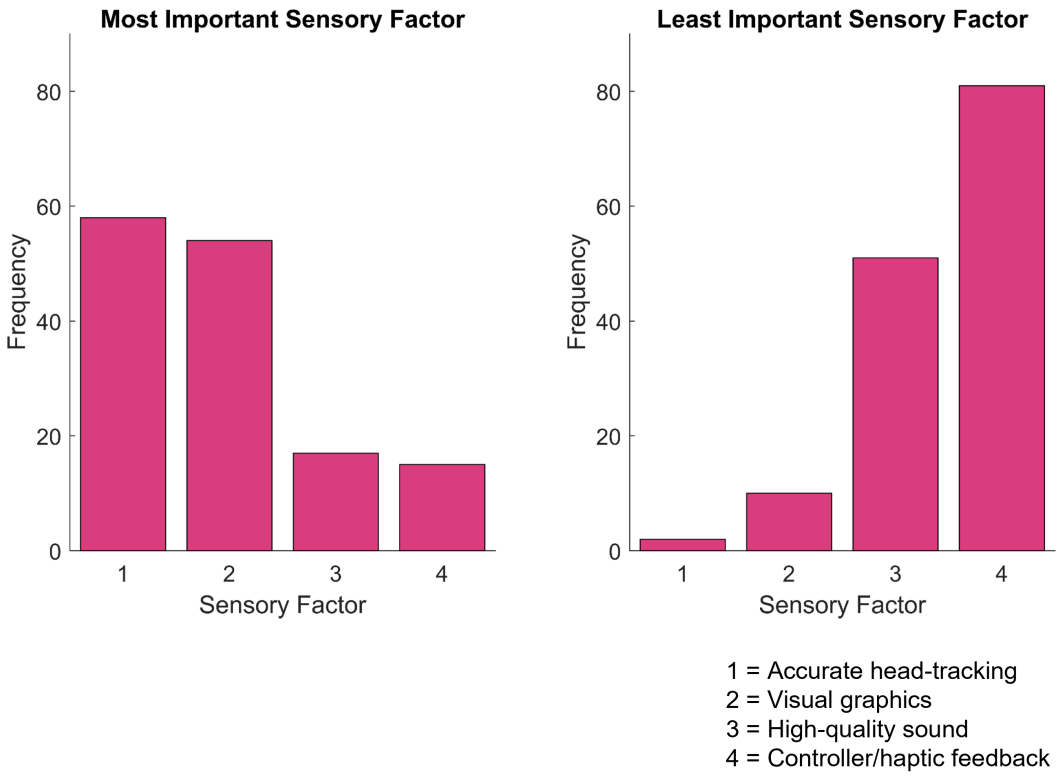


Figure 6.12: Frequency of the sensory factor rated most and least important for a VR device.

the technology, and as shown by the responses of an earlier question asking about the morbidity of motion sickness, are common. Relating back to many of the acceptance models, an individual's experience has an impact on the chances of them adopting behaviour which may explain why so many participants here reported to having tried VR but infrequently (*Figure 6.7*) and not recently (*Figure 6.8*).

Theme 2, where most of the terms have a beta below 0.03, is comprised of arguably unrelated terms. The top three terms ('data', 'safety', 'health') suggest a cohesive theme of user privacy and safety, which was a recurring theme within free-text responses to other questions, as can be seen in *Figures 6.14* and *6.15*. Additionally, Theme 3 has disparate terms, although by including all ten of the items the topic of immersion and presence can be linked, where the awareness of one's physical surroundings is impaired by having a headset on. The top terms here are neutral in sentiment, and all terms have a negligible beta.

Theme 4 has the highest beta correlations out of all of the detected themes. In-

terestingly, the theme becomes more apparent with the lesser correlated terms ('area', 'health', 'surroundings', 'protection'). Having a safe environment to use VR is a sensible idea and is often recommended within VR software during gameplay. Many accidents which have been publicised online (FailArmy, 2021) could have easily been avoided by ensuring there is enough open space around the individual to engage in the virtual environment without hitting or bumping into something or someone.

Participants were asked about any concerns they thought should be taken into consideration both when using VR technology (from the perspective of a customer), and creating the technology (from the perspective of a manufacturer). Safety precautions ('health', 'safe', 'safety', 'sickness') appeared 41 times in the free-text response to precautions for users (*Figure 6.14 (a)*), however this theme was not the most common when participants envisaged the concerns of the manufacturers. From this perspective, the most prevalent theme for concerns surrounded comfort ('weight', 'heavy', 'light', 'comfortable', 'comfort' frequency = 36), as can be seen in *Figure 6.14 (b)*.

The word 'data' appears in both clouds in *Figure 6.14*, showing the concern individuals have over the data that is collected by VR devices. General Data Protection Regulation (GDPR) is a regulation on data protection and privacy in the European Union and the European Economic Area, it regulates how consumer data is collected, stored, deleted, transferred and used (European Parliament And Council Of The European Union, 2016). VR equipment collects new types of data compared to playing a non-VR game or watching non-VR content, such as the potential for detailed tracking of the user's behaviour and movement in their own home, which may be the cause for concern among the public.

Related to this, direct questions were asked about opinions on the types of data that is collected from users of both VR and mobile devices. *Figure 6.15* presents sentiment clouds for the free-text responses to these two questions: there were fewer common words appearing in the question about VR devices which is why there are fewer words in the

cloud.

Participants had many overlapping views on how data is collected from mobiles, and within the transcripts it is apparent that the public are aware that lots of information about them is continuously being collected: *"It makes me uncomfortable to even think about it"*. However, they are also aware that there is little they can do to prevent it (*"I have a concern for my data security, data ownership and usage. Typically, I'll try and restrict the data that is accessed, although I am aware that is somewhat a losing battle"*) and they do not understand what their data is being used for (*"I get a strong feeling that often my data is used without me knowing what it is used for. I don't trust it but also don't see much way around it"*). Some individuals were actively taking a stand against this: *"I would delete a game if it didn't work because it wants access to my contacts"*, although past experiences have enlightened the public and cast a negative light on the data industry: *"Data privacy is a myth. Cambridge Analytica proved that. Our data can be weaponised against us to persuade us into voting for candidates, buying things online etc"*. Contrary to this, indifference to the issue did appear: *"I don't care, I have nothing to hide"* however this was an uncommon theme. More commonly was the idea of reimbursement for the data being taken: *"I've given up most if not all of my data for free! I should be getting paid for it"*, which has been put into practise with concepts such as Google Opinion Rewards (Google, 2021) although this is for the data participants willingly give up with full consent.

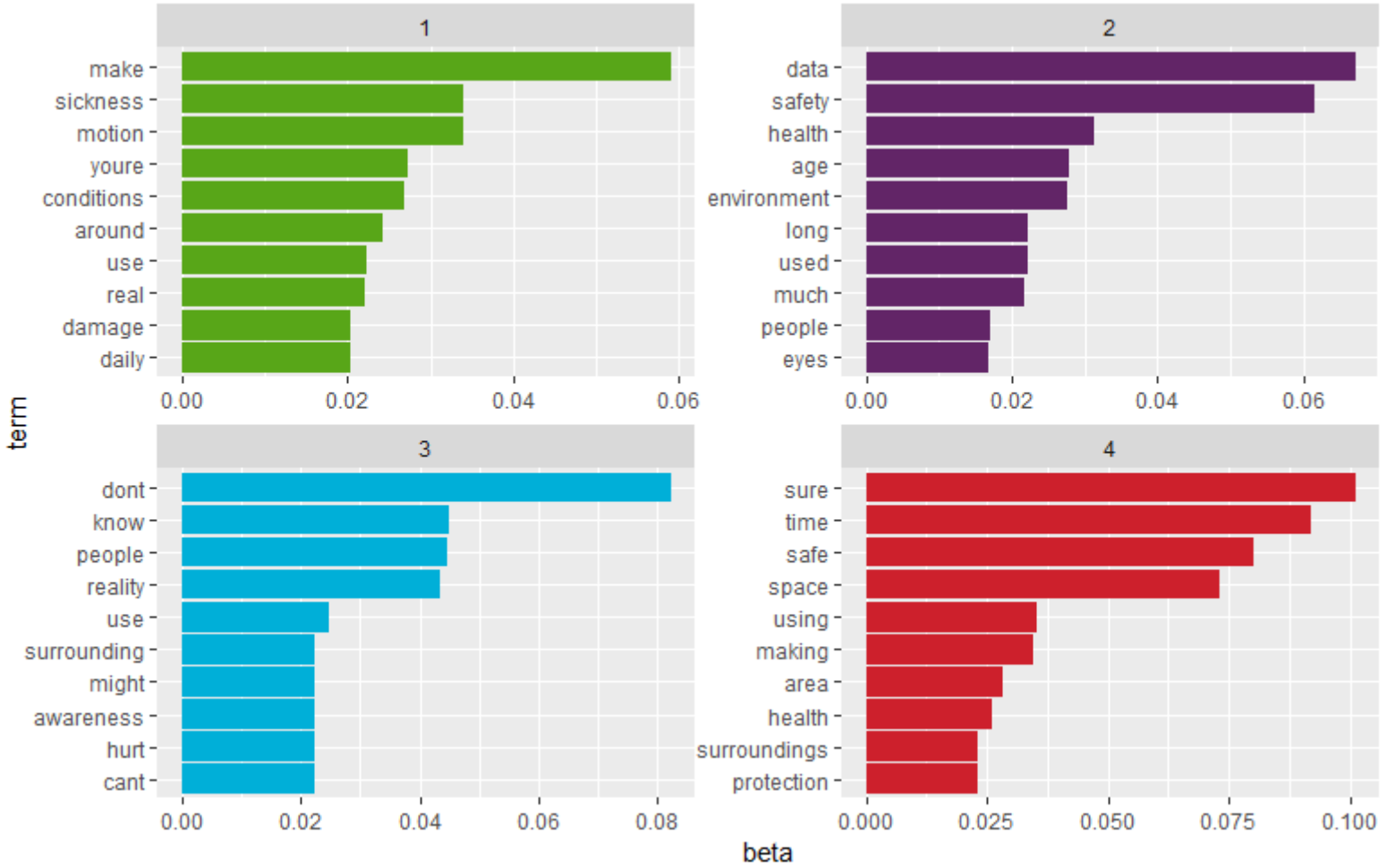


Figure 6.13: Topical analysis showing four themes detected by the CTM algorithm and the beta weighting of the 10 most frequent terms in response to the question 'what precautions do you think need to be taken when using VR?'. Topic *1* was interpreted to surround physiological risks; *2* privacy; *3* disconnect of virtual aspects from physical space; *4* secure environment.

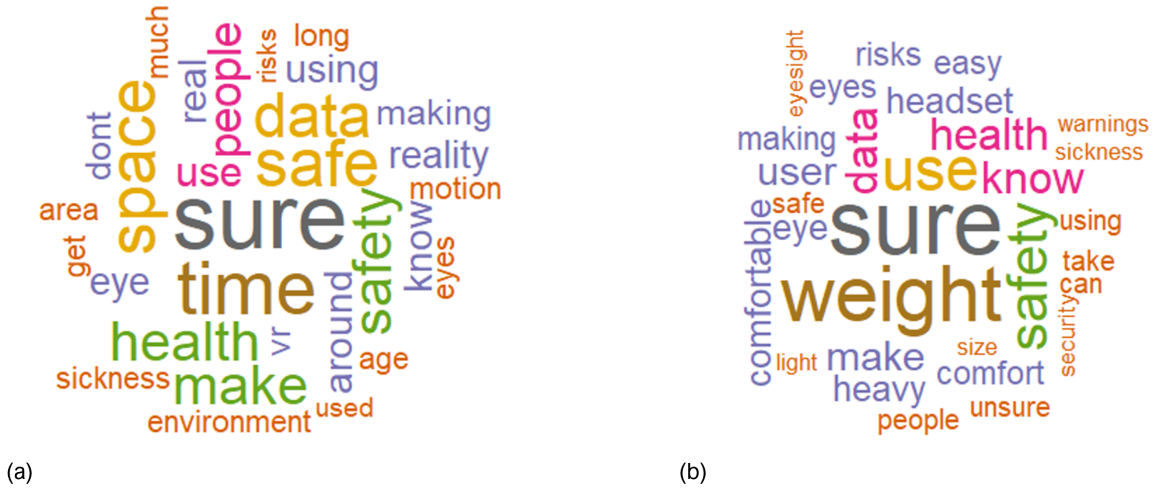


Figure 6.14: Word cloud for the free-text responses to the question 'what precautions do you think need to be taken when ...'; (a) prompt 'using VR?'; (b) prompt: 'creating a new VR headset?'. Larger words indicate higher frequency.

The responses to the question pertaining to data collection through VR devices followed the same ideology, not being on board with the collection of data in the first place (*"I would really rather that my data isn't collected, it's an invasion of privacy"*), while also wanting to understand how the information will be handled (*"I need to know what it is, what it is used for and whether it will be anonymous"*). As the technology is still relatively new and developments are still underway, individuals are sympathetic to this: *"If it is for optimizing the technology then it is ok, if it is for the sakes of profiting then I am strongly against"* although there may be a disconnect between what the manufacturers would consider a valid use (*"I don't mind them using my data only if it's for a valid reason"*). Nonetheless, individuals feel some sort of compensation should be offered for the personal data that is accumulated (*"I'd like to know what data, how its used, and get paid for if I give it up"*).

The word 'like' appears frequently in the responses to both questions: in the free-text responses this word is used to express emotion (*"I do not like that they can use and see our data"*) as well as a filler term (*"Seems like a lie to try and get data"*). In addition, 'better' is frequent, although only used to express interest (*"sometimes needed for better gameplay"; "the lesser the better"*).

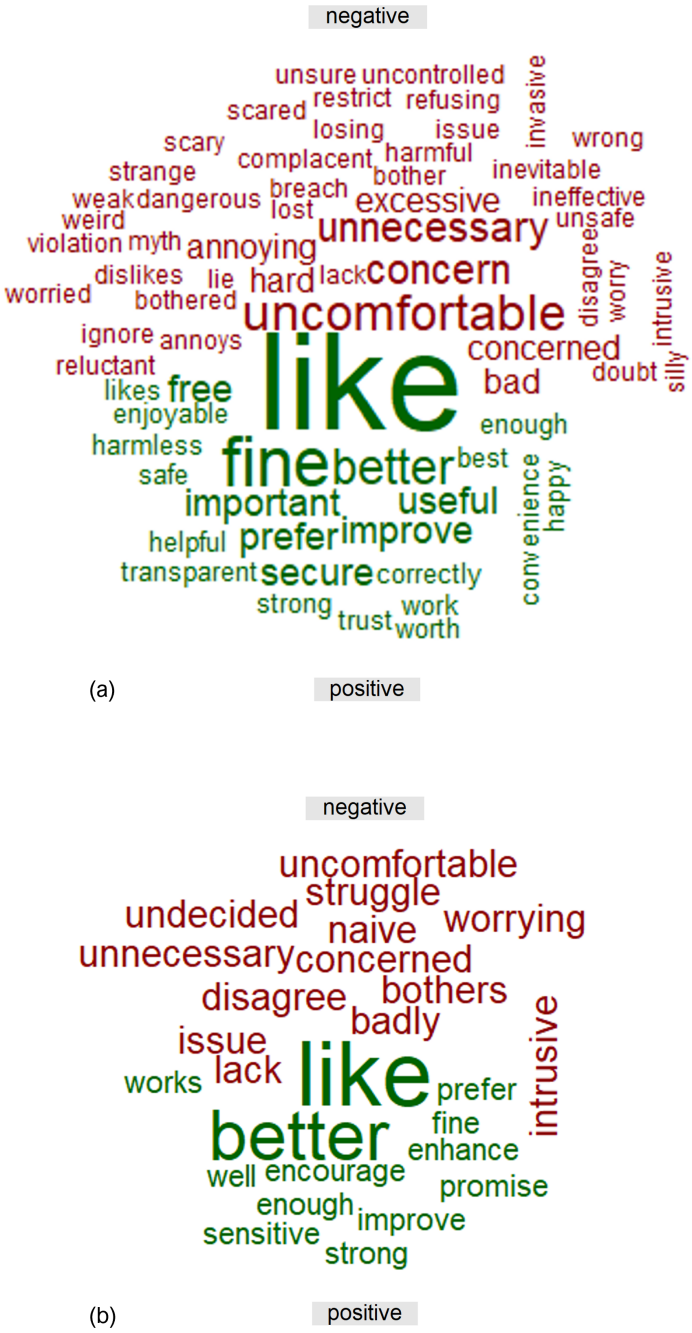


Figure 6.15: Sentiment analysis clouds for the free-text responses to the question 'Do you have any opinions on the data that is collected on you, from ...?'; (a) prompt: your mobile phone; (b) prompt: a VR headset. Larger words indicate higher frequency, sorted into positive and negative categories.

6.3.4 Opinions

Those participants who reported not having tried VR were asked why; *Figure 6.6* is a word cloud of the free-text responses. The word which appeared most frequently was 'opportunity': a positive connotation cannot be inferred from this, this does not mean that these people want to try it but they haven't had the chance to. It can only be concluded that there is still a lack of availability of the equipment to those who are not actively involved in the technology.

All participants were asked what they associated VR with. Six negatively associated words appeared multiple times in individual's responses: *Figure 6.16* shows how the cost of the equipment, the fact that what is displayed in the technology is not real, and the idea of it being a 'threat' are common associations. These link to many of the aforementioned acceptance models, as well as the responses for other, much related questions: *Figure 6.14 (b)* and *Table 6.1*.

When asked to rate the appropriateness of different factors which can be found in current technologies, the consensus tended to be similar for both mobile and VR devices. The three question prompts and the most common participant responses can be seen in *Figure 6.17*, where similar levels of appropriateness for facial recognition, eye-tracking, and content personalisation are found between the two distinct types of technology. The results of the Wilcoxon signed rank tests are presented in *Table 6.2*. This shows that participants felt that facial recognition and eye-tracking were appropriate in mobile devices and VR. They were overall neutral on the subject of content personalisation on both platforms. Interestingly, content personalisation had mixed reports, with enough variance for the test to report no significant direction in the unison of responses. Content personalisation may include personalised advertisements, or content recommendations; a survey by Segment (2017) found that on average, 71 % of consumers expressed frustration when their shopping experience was impersonal, which is somewhat in opposition

Table 6.2: Output from the one-sample Wilcoxon signed rank test on the 5-point likert style question 'what are your thoughts on ...'.

Prompt	V	p	Direction
Facial recognition on mobile devices	6057	<0.001	Positive
Facial recognition on virtual reality devices	4188.5	<0.001	Positive
Eye-tracking on mobile devices	8272.5	<0.001	Positive
Eye-tracking on virtual reality devices	7308.5	<0.001	Positive
Content personalisation on mobile devices	3615.5	0.899	Neutral
Content personalisation on virtual reality devices	4348.5	0.003	Neutral

to what was found here. Of course, the survey was pertaining directly to shopping which is why the difference may be so apparent, whereas in the present questionnaire the question was regarding VR which has many uses not limited to shopping.

Participants were asked to report their thoughts on the notion that VR may become a daily commodity. *Figure 6.19* displays a sentiment analysis conducted on the free-text responses to this question. Opposing words appear in both sides of the cloud, for example 'useful' and 'useless'. The size of the font depicts the number of times the word appeared in the answers, here the frequency of 'useful' was 11, and 'useless' was two. In this cloud, 'useful' had a higher frequency than 'fun', which aligns with the TAM but is the opposite for the VR-HAM. In addition, 57.64 % of participants could see VR becoming a daily commodity in the next few years.

Participants were asked what ways they thought VR could benefit them personally. Occurring 19 times was the word 'entertain'; related words such as 'game' (n = 10), 'gaming' (n = 12), 'fun' (n = 17) mean that the theme of recreation was the most common in the responses to this question. This coincides with the forced choice options from the question shown in *Figure 6.9* where recreation was the most chosen answer to what VR was being utilised for by current users.

The word 'stress' appeared five times in the responses to what VR was being used for, in contexts such as *"take my mind off of daily stressors in real world"*. Having an



Figure 6.16: Sentiment analysis conducted on the free-text responses to the question 'what do you associate VR with?'. Larger words indicate higher frequency, sorted into positive and negative categories.

immersive environment that causes one to forget about the physical world is a useful tool, not only for reducing stress but also for medical interventions: *"help with fears and PTSD, maybe even just from escaping this world"*. Especially in the current climate, where the unprecedented events are accumulating, it is no wonder this form of distraction is desired by so many individuals.

In this context, experiencing places or events was a common answer, one individual elaborated: *"because of the pandemic I could not come for a campus tour but having that campus tour experience on VR could be the next best thing"*. This could indeed be used as a great tool in allowing individuals to attend a university that may be hard to reach for an open day, or giving potential customers a taste of a service a business has to offer.

Other themes, less common but appearing numerous times, include using the technology for exercise, training, and research. There are currently software programs being used in these domains, and it is positive that the public acknowledge these as beneficial uses. On the other hand, using VR as a release from boredom appeared three times.

One free-text question asked participants to explain what improvements they thought should be made to VR devices. Many individuals gave explanations for their suggestions,

however the majority gave a few word descriptors. Descriptors surrounding the hardware (weight n = 22, size n = 5, wires n = 4) appeared the most times, with all who mentioned this requesting the device to be made smaller, lighter, and without cables. The weight and bulkiness of devices was noted in *Figure 6.17*, so it is reaffirming to have the same themes appear without prompts.

Having a reduction in the cable length can also be linked to safety, with the removal of cables causing a reduction in tripping hazards. One participant added: *"incorporating some form of physical apparatus that can truly immerse yourself in a new virtual world without major risk of injury"*. The addition of physical apparatus to reinforce the virtual guardian system, which is present in many devices, is suggested here to enhance the immersive experience, however it might work in the opposite way as a grounding feature to the physical world.

The inclusion of a break announcement, when the user has been involved in a virtual environment for an extended period of time, was an improvement that one individual suggested. This would aid in reducing some of the other physiological side effects of the technology that has arisen in other questions, such as eye fatigue.

One question required participants to rank some of the top VR manufacturers in order of best to worst (*Figure 6.18*). The mode response for the number one ranked brand was HTC, who have devices ranging in capabilities and prices suitable for novice and dedicated users. At the time of writing, headsets ranged from £499 to £1,499 (Vive, 2021b). With a price range this broad the fact that there are intended target audiences for the different devices is obvious.

The brand ranked lowest was Valve: again, at the time of writing this brand only had one available device (Steam, 2021), which could be a determining factor as to why this was ranked worst and HTC was ranked best.

As a follow-up question it was asked what participants based their answer to the previous question on. This was a closed question, with options to select either having

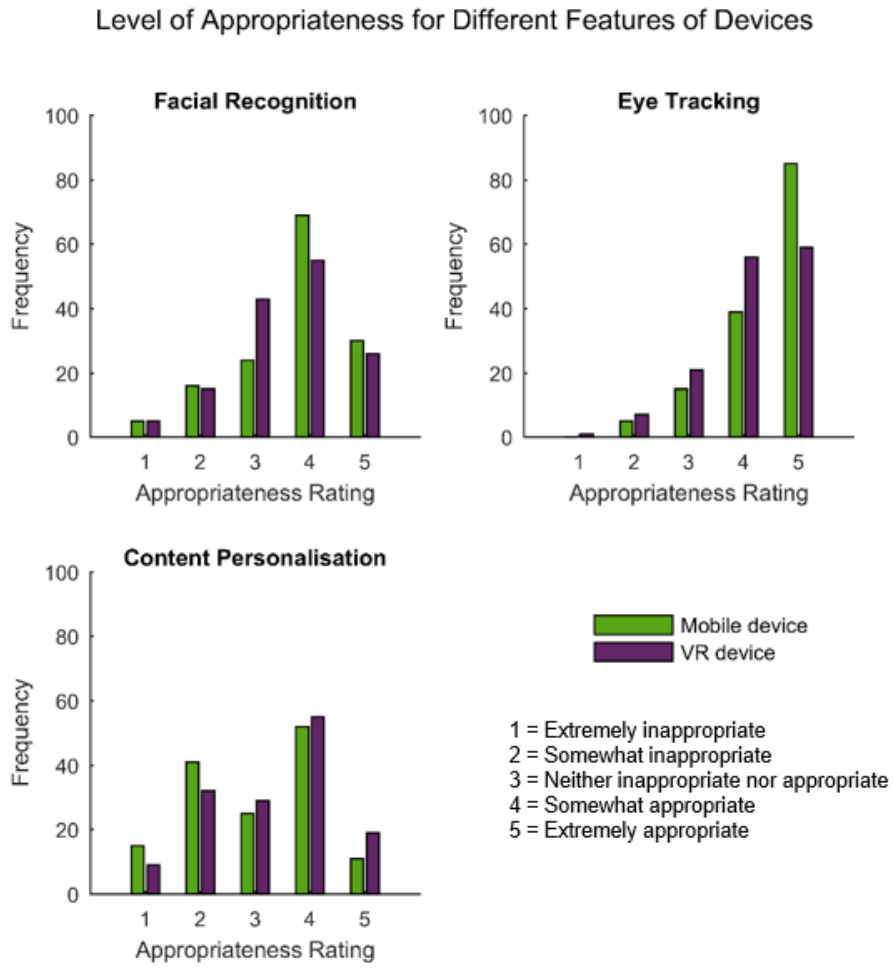


Figure 6.17: Ranked level for how appropriate each type of data collection is perceived to be, for both mobile and VR devices.

'knowledge on the bands generally' (n = 63), the individual's 'own experience' (n = 31), or 'other' (n = 45). Many of those who selected 'other' elaborated in a free-text option, by stating that their opinion of the brands ranking was based on a combination of both prior experience and general knowledge. A number of individuals declared that they had only heard of one or two of the brands and therefore ranked those first, others reported to not having any knowledge on any of them so the order was uninformed and random (n = 22).

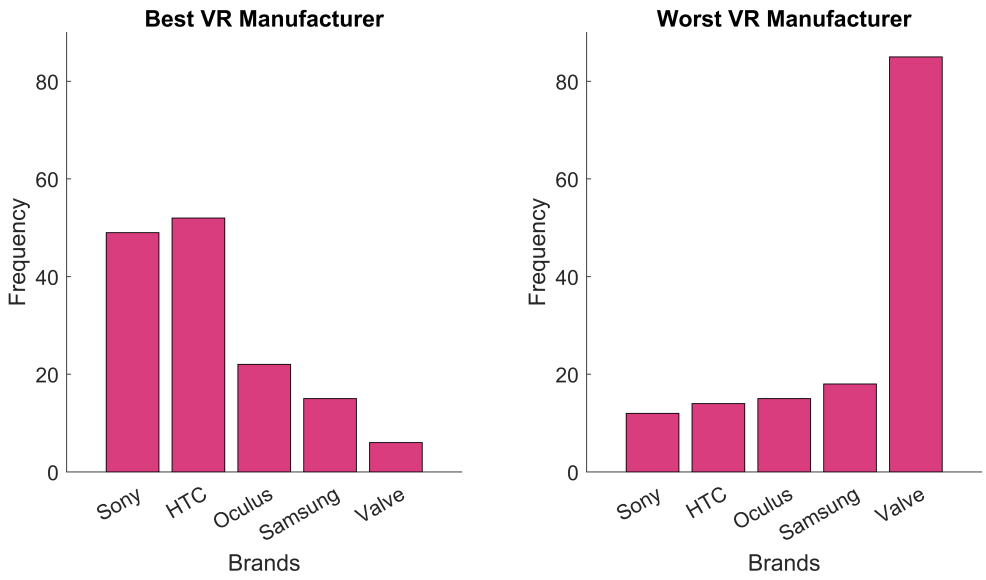


Figure 6.18: Frequency for the of the VR brands rated best and worst.

6.4 Conclusions

6.4.1 Themes Found

It has been suggested that direct product experience is the best way for consumers to evaluate a product (Gartner, 1994; Hyun & O'Keefe, 2012; Klein, 2003). Due to this there were branches in the questionnaire that appeared only to those who already had familiarity with VR. In addition, some of the analysis was conducted to collate the opinions of only users of the technology.

The topic of health and safety appeared in the responses of numerous different questions (*Figures 6.13, 6.14, 6.19*). Individuals have shown concern over the side effects that are common when using VR technology, such as motion sickness and eye fatigue. Related factors within this topic include tripping hazards, the amount of space required to use the technology, and reduced awareness of one's surroundings.

Components of the hardware, such as weight and form factor, arose in multiple responses: the term 'comfortable' in *Figure 6.14 (b)*, 'uncomfortable' in *Figure 6.15 (b)*; 'bulky' was agreed upon by the participants as an appropriate term to describe VR

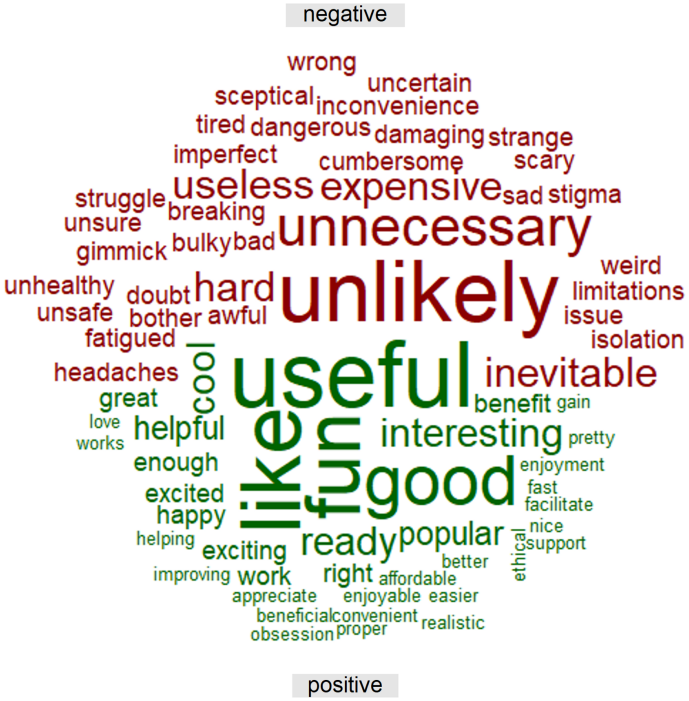


Figure 6.19: Sentiment analysis conducted on the free-text responses to the question 'what are your thoughts on VR becoming a daily commodity?'. Larger words indicate higher frequency, sorted into positive and negative categories.

technology in *Figure 6.11*. Likely a result of the hardware composition, the pricing was also a notable theme within the responses of this questionnaire: in *Figure 6.10* the price was found to be the second most important feature rated for the technology, and the weight was ranked third.

A concerning issue of privacy was unsurprisingly apparent here. Participants feared for the way their personal data is used (potentially against them) and without their informed consent. Development of eye-tracking in VR was accepted as being appropriate, however this raises the question of whether iris recognition will become another area for development, as it has been in mobile phones. Will headsets store unique information about user's iris texture? At the moment, facial recognition was voted to be somewhat appropriate for VR devices, although it is not inconceivable that iris recognition is on the horizon. In the same regard, content personalisation had mixed reviews for its place in VR; this could include the presentation of advertisements in the headsets.

6.4.2 Implications for Manufacturers

Even though VR technology has been around for decades, it is currently more available and usable than it has ever been. Singh et al. (2020) stated that VR has played a vital role during the Covid-19 pandemic, allowing for medical marketing, disease awareness, medical training and learning. If more applications along these lines can be developed then the target audience for the equipment would widen, as would the use and public reach. Throughout the pandemic, the importance of physical activity has become apparent: over half of the population reported wanting to do more physical activity once the pandemic restrictions eased (Sport England, 2021), it is for this reason that applications of the technology should be implemented into the domain of sports. VR software is currently being used in training for football (Shimi et al., 2021), volleyball (Du & Li, 2021), and perception-action tasks for predicting opponent's behaviour with rugby and handball (Bideau et al., 2009). Considering that the bulkiness of the headsets and price are an issue for the consumers of VR devices, these factors need to be modified by the manufactures in order to increase the likelihood that medical professionals or sports teams adopt them in their training practises.

Haptic feedback was not considered to be an important sensory feature of VR by the participants of this study. The majority of efforts from manufacturers should therefore be placed onto enhancing the head-tracking and developing the available content, based in the results found here. This, alongside reducing the weight and bulkiness of the headset itself, would go a long way in pleasing potential and existing consumers. Other hardware obstacles include the focal distance and accommodation conflicts, one individual pointed out that efforts need to *"compensate for eye focus difficulties, short or long-sightedness"*.

The sense of being immersed in the virtual environment is a central aspiration that is being researched. In one of the free-text questions, an individual stated that *"single player games appear to be a lot more immersive"*: compared to what was not stated,

however if compared to multiplayer games then this raises an intriguing question as to what components are different in solo versus multiplayer games that inhibit the level of immersion experienced by players. As the only difference between these types of games is a result of the software and not hardware, games developers should invest in figuring this out if the intent is for the consumer to have an immersive cooperative gaming experience with others.

6.4.3 Comparison with Manufacturer Responses

In a survey by Perkins Coie (2020), industries who were involved in VR/AR/MR technology provided responses on the factors that are important for the adoption of these technologies. A total of 191 professionals from established technology companies (47 %), start-ups (19 %), consultants (16 %), investors (12 %), and policy makers (5 %) took part in the survey in early 2020. The following section compares the responses found in the current study, to those reported in that survey.

One question in the survey asked what the biggest obstacle to mass adoption of AR and VR was. The most common answer for VR was content offerings, followed by user experience. For AR user experience was judged the most important factor, followed by content. In the description used in this study, user experience included features such as the bulkiness and technical issues; as seen in *Table 6.1* participants did agree that the VR headsets were somewhat bulky, and had an ambivalence to whether or not they would own a headset in the future. This, in combination with *Figure 6.10* is in agreement with the industry survey. However, purchase price was ranked second most important here, whereas only 11 % of manufacturers ranked the cost to consumers as important. In another question, manufacturers reported that a reduction in the size of the devices would be the top improvement to VR software in the following two years which will come as a relief to many individuals in this study who commented on the form factor of the current headsets.

In response to the question asking the manufacturers to identify what their biggest concern to their own organisation was, 61 % stated consumer privacy/data security in 2019, which reduced to 49 % in 2020. It is reassuring that the companies are aware of this being an issue, as this also appeared in the free-text responses in the current study when participants were asked what precautions the manufacturers need to take into consideration. A word cloud is shown for the responses in *Figure 6.14 (b)* related to this question. Unsurprisingly, data protection did emerge in the responses, both in the responses for what manufacturers need to take into consideration and also for users of the devices. To add to this, the theme also appeared in the topical analysis that was conducted on the same question. Yet, the reduction in manufacturers reporting this as an issue from 2019 to 2020 is concerning, especially given the continual rise in GDPR violations, increasing 113.5 % from 2020 to 2021 (Baltrusaitis, 2021). Just over half of the manufacturers (54 %) stated that they had updated their privacy policies and disclosures regarding consumer data, however less than half (48 %) stated that they were limiting the amount of personal information from consumers that are collected/shared/used. The manufacturers are not currently taking the steps desired by consumers to overcome these privacy issues, instead they are covering their own backs in legal loopholes of burying information in the terms and conditions that are not read by consumers (Steinfeld, 2016).

A total of 47 % of manufacturers stated that they would earn from having product placement within the immersive experience of VR devices. This might be beneficial for those in *Figure 6.17* who deemed content personalisation as somewhat appropriate inside VR, if they visualised getting personalised recommendations for other games they might enjoy, but not advertisements for non-related products or services being offered in the physical world.

Improvements to the technology were asked to be suggested in the current study, where alterations to the hardware appeared to be most agreed upon. The manufacturers were asked a similar question, to state what improvements would make the greatest im-

pact for the consumers: regrettably, this was not an area that ranked even top five. Solutions for application experiences, bugs, compatibility, connectivity, and cyber-security were all detailed as being more important for the manufacturers when taking the perspective of the consumers. This bifurcation reinforces the need for the creators to invest more time into user research to understand their requirements.

In the current study, participants were asked to rank some VR brands from best to worst (*Figure 6.18*), the Valve Index equipment was rated the worst, which coincides with the number of variations of the device that the company has to offer (only one version on the market). The manufacture's survey included questioning about the most common concerns from investing in immersive technology, and the top voted reason was a lack of an established market for the technology. Valve Corporation are an established gaming developer, with little to no involvement in other immersive technologies. This may be the basis for why participants voted it lower, as there were no prior attempts to enter the market with this equipment, compared to the other brands with multiple, successful immersive products.

Less than half of the participants here thought that VR would become a daily commodity in the next few years; more manufacturers reported immersive technologies to become mainstream in the two to five years after the survey than any other time period.

It was reported that 76 % of manufacturers anticipated that the AR market would surpass the VR market, mainly due to the cost and accessibility of AR compared to VR. Most expected this to occur within three to five years after the survey was conducted. In the current study, only 36.73 % of participants had tried AR compared to 67.37 % who had tried VR. It is a naive assumption that the majority of individuals will contribute to the success of AR over VR given that most have not tried it. With developments in the technology, marketing, and accessibility, this may change in the future, however based on acceptance models, it is safe to assume that they have little intent to try the technology, or even curiosity in it for the time being.

6.4.4 Summary

This questionnaire explored the opinions and demographics of those who use virtual reality equipment, as well as delving into the reasons of *why not* for those who did not use the technology. It was revealed that the most common use of VR was recreation, and a majority of individuals who had tried VR did not use the technology often. Although commonly associated with having fun, the responses here uncovered some negative factors which may be an important focus to resolve for manufacturers. The most important factors uncovered here were the bulkiness of the headsets themselves, and the data security issues which these devices enable.

Chapter 7

General Discussion

In academic science there is a tendency to use robust, albeit dated, technology. There is a greater level of advancements in the set-ups that are aimed at general consumers rather than researchers, however the virtual reality (VR) equipment that is readily available to these consumers is yet to be validated with certainty that it can deliver the same amount of reliability that the purpose-made equipment has provided. This thesis has provided evidence that VR is able to replicate the same perceptual outcomes as esteemed research has found using specialist equipment, in the domain of academic science. In addition to identifying the extent to which VR has potential, this thesis examined the contributing factors to which perception can be enhanced/diminished through the use of alternative programming styles that can be used to create such environments. An explanation into how these perceptual experiments can be coded in a free-to-use games engine has also been presented. This final dissuasion chapter will review the findings of each experimental chapter, while cross-examining the outcomes with similar work, as well as putting forward ideas for continuation into this research area.

7.1 Overall Findings

7.1.1 Recap of Thesis Research

Each chapter within this thesis examined a different topic surrounding the way virtual stimuli presented in VR is perceived. Below is a brief overview of each chapter outlining the aims, findings, and impact of the research.

Chapter 2 provided a brief overview of quantitative experiments and methods, and the contribution that VR can make in expanding the range of methods available to researchers. Its primary contribution was a walk-through tutorial for creating a basic psychophysical experiment in VR using Unreal Engine. Justification has been given as to why this program is effective for use in scientific studies, and discussion surrounding why different experimental set-ups may be more appropriate than others.

In *Chapter 3*, shape and size constancy was assessed for stimuli presented in two consumer VR systems, using software developed with the same platorm. Participants wore either the HTC Vive or the Oculus Rift and adjusted the size of a virtual stimulus to match the size and depth of a physical stimulus they were able to refer to haptically. The judgements participants made allowed for an indirect measure of their perception of the egocentric, virtual distance to the stimuli. The data show under-constancy and are consistent with research which used carefully-calibrated psychophysical techniques (Brenner & van Damme, 1999; Creem-Regehr et al., 2015; Mon-Williams et al., 2000). There was no difference in the degree of constancy found in the two headsets. Here, it was concluded that consumer VR headsets provide a sufficiently high degree of accuracy in distance perception, to allow them to be used confidently in future experimental vision science, and other research applications, in psychology.

Chapter 4 aimed to test how different visual cues contribute to the accuracy and precision of the perception of distance through two paradigms. In the first task, par-

ticipants positioned two spheres so as to divide the egocentric distance to a reference stimulus (presented between 3 and 11 m) into three equal thirds. In the second and third tasks, participants set the size of a sphere, presented at the same distances and at eye-height, to match that of a hand-held football. Each task was performed in four environments varying in the available cues. The `Sparse` environment presented only a floor plane; linear perspective was added by presenting walls and a ceiling to create the `Perspective` environment. The `Texture` scene added bricks and a floor pattern; the `Cluttered` environment had numerous objects added in. Accuracy was measured by identifying systematic biases in responses, and precision as the standard deviation of these responses. While there was no evidence of non-linear compression of space, participants did tend to under-estimate distance linearly, but this bias was reduced with the addition of each cue. The addition of binocular cues, when rich pictorial cues were already available, reduced both the bias and variability of estimates. These results show that linear perspective and binocular cues, in particular, improve the accuracy and precision of distance estimates in VR across a range of distances typical of many indoor environments.

The experiment in *Chapter 5* aimed to identify whether a target's shadow is used as a cue to distance. This was measured by assessing the precision and directional bias of distance estimates. Six lighting conditions were created and presented in VR for participants to perform a perceptual matching task. The task was repeated in a cluttered and sparse environment, where the number of cast shadows (and their placement) varied. No significant difference was found between the sparse and cluttered environments. This could be explained by individual differences between participants: some utilising the objects as anchors to aid them, while others finding them distracting. Having an ambient light source produced the most variable and inaccurate estimates of distance, however the difference between performance in this lighting condition and the next worst was not significant. Lighting positioned behind a target reduced the under-estimation of distances perceived, and the most precise estimates were as a result of a light source

being positioned directly above the target.

An exploratory study was presented in *Chapter 6* where both qualitative and quantitative data were obtained about people's experience of and attitudes towards virtual reality. This study collected data from both experienced users and novices about their perceptions of VR equipment, the concerns surrounding data collected by the equipment, and facets needed to be taken into consideration for future developments. It was revealed that data security is still an issue which has yet to be overcome. Individuals were aware of the risks associated with the newer technology, in terms of privacy and also potential physiological side effects. The opinions surrounding precautions and important VR features from the public did not differ drastically from the opinions of manufacturers, however this study provides an in-depth perspective into the improvements desired by the public.

7.1.2 The Value of Using Virtual Reality

The types of headsets readily available to consumers now have high-quality specifications, meaning researchers are able to switch to these devices and avoid paying premiums for research-grade equipment. This does come at some costs: in current consumer devices it is not possible to alter the focal distance. Specific multi-focal displays which have been created and have been found to provide near-correct focal cues (Akeley et al., 2004; Watt et al., 2005b) however these are currently bulky and not yet suitable for consumer-use. Not limited to physiological side effects such as visual discomfort and eye fatigue (Shibata et al., 2011a, 2011b), the conflicts arising from focal distance inconsistencies also contribute to distortions in depth (Girshick et al., 2004; Watt et al., 2005a). The studies presented in this thesis did find some biases in perceived depth, and these conflicting focus cues could have contributed to these biases to some degree.

It is also easy to sidestep the limitations which arise from the programming software that has traditionally been used for programming experiments: instead of learning

coding languages to the degree necessary to create graphics with high levels of realism and responsiveness, it is now possible to use games engines to create realistic and controlled experiments, as presented here. Timing is a limiting factor however, especially when creating experiments in games engines such as Unreal Engine. Psychophysical experiments often require a higher refresh rate to capture accurate response times than what are used during game play; the current version of the games engine (UE4) maxes out before accurate timings to the millisecond can be achieved. With these engines the stimuli can be precisely manipulated in terms of the actual cues visible to an observer in any VR device. While able to maintain careful control, VR provides a mobile and dynamic viewing experience of a controlled environment.

There is debate over which sources of information should be included and rendered correctly in a VR simulation, and which can be safely ignored or presented in a degraded form (Scarfe & Glennerster, 2019). The experiments in this thesis have aimed to explore the options for these types of visual cues in a naturalistic viewing environment. In the experiments presented here, an unconstrained viewing experience was provided, although observers were not required to actively explore the environment. Movement can provide important information for the observer in terms of how the images of the scene adjust in response to movement of the observer, which is not available to a static observer (Koenderink & van Doorn, 1987). The Gibsonian perspective of psychology suggests the importance of context, in that perception occurs only when purposeful observers actively seek stimulation; this theory is better tested in VR compared to passive displays, as observer movements can be monitored, recorded, and analysed. Models developed to explain behaviour in moving observers are different from those that are applicable to an experiment in which the participant is fixed to a bite bar (Scarfe & Glennerster, 2015) or observing a passive display, because in a perceptually static environment, void of engagement or under constrained viewing conditions, the viewing experience is subjectively different from natural viewing. Development in the technology has led to VR models which are not tethered by wires (Oculus, 2021) or have wireless adaptors (Vive,

2021c), so as to encourage movement as part of a more active experience (Thomas, 1999). This coincides with the notion that vision should be treated as an exploratory activity (O'regan & Noë, 2001).

It is common to undergo careful calibration of displays, and control observer movement, in traditional psychophysics; this contrasts with the use of consumer hardware, and freely moving observers. A set of experiments by Backus et al. (1999) aimed to identify how the visual system determines orientation of isolated surface rotating around vertical axis. Through five experiments, viewing conditions were manipulated for the maximum of four participants who took part. It was concluded that size ratios of a stimulus and senses eye position are weighted differently by the visual system, depending on the information observable. Although the number of participants is contentious, individual calibration of the monitors were undertaken for every user, along with a bite bar to ensure the intended stimuli were observed correctly. While this approach is effective at testing the mechanisms the visual system uses during passive viewing, these results are not representative of how the visual system incorporates cues from the other senses, or the information retrieved by exploring an environment dynamically. While it is important to have a calibrated system to avoid discomfort when fusing stereo images (Mon-Williams et al., 1995; Wann et al., 1995), it could be argued that this carefully calibrated set-up is not representative of the way consumers would set-up the equipment. This has led to manufacturers developing improved built-in calibration methods for head-mounted display set-ups.

Gilson et al. (2008) presented a quick way to calibrate an augmented reality display which resulted in few re-projection errors and no estimates on the geometry of the device. This has shown to be useful, however consumer-grade devices do not have adjustable calibration anywhere near this extent. There are brief calibration adjustments which a consumer can make, however this leaves questions about how accurate these processes are. In consumer VR there is the assumption that all calibrations are taken care of behind the scenes and is correct for accurate viewing. A study by Niehorster et al.

(2017) tested the tracking of the HTC Vive to find that the reference plane used within the headset was actually tilted away from the physical ground plane. This resulted in errors of tracked height, as well as roll and pitch measurements, in addition to the tilt of the reference plane altering when the headset lost and regained tracking. Due to these offsets, the authors concluded that this device would not be suitable for research where self-motion through a virtual environment is required. The assumption that consumer headsets are appropriate for all types of scientific research should therefore be questioned, as the applicability will be task-dependent and device-dependent.

Individual differences, such as the distance between one's eyes (IPD), is a factor in which one calibration or version of hardware cannot cater for everyone. Binocular information has been found to impact perceived quality and magnitude of depth in stereoscopic displays (Hibbard et al., 2017a) so any mismatch between an observer's IPD and the simulated IPD of the viewing device should therefore result in errors and a subjectively different experience compared to natural viewing (Scarfe & Glennerster, 2015). A paper by Hibbard et al. (2020) explains the way vergence and IPD are used as distance cues by the visual system, and discuss the issues that are experienced by an observer as a result of mismatched observer IPD and simulated IPD within a VR device. The authors concluded that, from testing in three different consumer devices, the main problems include visual discomfort, off-centre viewing, divergent fixation, and biases in depth perception. Evidence from *Chapter 6* shows that consumers do not tend to set their devices to match their own IPD, likely because they do not know the importance of this. Throughout all of the studies presented in this thesis, the IPD of headsets used was not individually calibrated to each observer. Statistically, the equations presented in Hibbard et al.'s article suggest mismatches are detrimental to both the subjective and objective viewing experience; although, the biases found throughout this thesis were comparable to carefully calibrated 3D set-ups, and participants reported to still enjoy using VR, even when this mismatch is present.

The experiments presented in this thesis have used consumer VR devices. There have been no direct comparisons using the same methods in physical space, or in any carefully calibrated 3D displays. Different display devices will vary in the graphics and specific visual cues available, and potentially even display properties. Although it is important to recognise other devices and physical-world viewing have these differences, the results found in this thesis are comparable to studies conducted in other environments such as these. One of the conclusions drawn here is that off-the-shelf VR is a viable alternative, while reproducing general trends of distance mis-estimations.

A useful aim for research is to combine real-world complexity with experimental control, which can be accomplished in VR. In using VR, environments are able to be created with scans from physical objects, and specific cues can be manipulated using software such as Unreal Engine. This provides the perfect ground truth, such as the previously specified 3D location of each point in the environment, which is typically not available in natural environments. Conditions are able to be repeated with specific adjustments, resulting in a high level of experimental control which is difficult to combine with naturalistic settings in any other circumstance. These factors boost the ecological validity of any scientific research conducted using these set-ups, which can therefore lead to more believable experiences, more efficient training simulations, and more generalisable findings of research.

7.2 Future Directions

7.2.1 Exploring the use of Unreal Engine 5

Unreal Engine 5 (UE5) was expected to become available in 2022 (L'Italien, 2021) and should allow for the creation of even more realistic virtual environments. Updates for this which enhance its capabilities over version 4 include Nanite geometry (causing there to be no need to bake details to the levels, no polygon count budgets, no draw count budgets, no polygon memory budgets) and Lumen dynamic global illuminisation (for

diffuse inter-reflection with indirect specular reflections capable in kilometre-wide environments down to the millimetre) (Epic Games, 2021). With every step that enhances the visual components and reduces artefacts, the ecological validity for the results of experiments in such environments increases.

If it can be shown that an experiment is able to elicit similar physiological responses, and sensory perception in VR as would have arisen in physical space, then the possibilities for using computer simulations opens up an even larger scope than there already is. Using a higher-powered developer programme such as UE5 could allow for enhanced training programmes for surgeons and pilots with more precise learning outcomes, due to the immersive on-site and risk-free training. Of course, scenarios such as these would require testing to ensure the VR environments are indeed eliciting lifelike physiological responses as well as perceptual ones.

7.2.2 Higher-level Influences on Perception

Top-down processing has been shown to affect perception (Bruner & Goodman, 1947; Gantman & Van Bavel, 2014; Sciuto & Hartley, 1963). In an influential study by Bhalla and Proffitt (1999), the authors found that participants wearing a heavy backpack perceived the gradient of a hill to be steeper than when no backpack was worn. This is just one of many applications of the idea that prior assumptions influence perception; with a spike in recent years of public interests over health and obesity, it would be topical to apply it to perception of food.

The proposed study would use VR as a method of presenting naturalistic food and non-food stimuli in a virtual environment. A difference in size perception of food has been found when participants are in differing mental sates (hungry versus satiated), however much of this research is conducted on 2D displayed in laboratory settings (Collier, 2018; Crutchfield et al., 2018). Moving this research over to the realm of VR will enhance the applicability of the findings due to the enhanced realism that comes with immersive devices and 3D rendering of naturalistic stimuli. With presentation distance

varying, participants would take part in the study on two occasions: once immediately after a meal, when they were satiated; and again when they had not eaten for a set amount of hours, when they were hungry. The order for these conditions to take place would be counterbalanced between participants, and a questionnaire would be distributed before each session to ensure two physiological hunger states were met. The task would be for participants to manipulate the height and width of a cross hair, so that it matched that of the (non)food item presented in each trial. Based on current findings in 2D (Brunstrom, 2011; Sakurai et al., 2013; Van Koningsbruggen et al., 2011), the hypothesis would be that the size estimates from the hungry condition would be less accurate; there would systematically be larger size settings compared to those from the satiated condition. In addition to this, it would be hypothesised that size estimates for the non-food stimuli would be more accurate than the food stimuli, regardless of the condition.

This novel experiment would present yet another application of VR in research, with a validated methodology that has been presented throughout this thesis.

7.2.3 Multi-sensory Cue Combination

Wallraven et al. (2014) examined whether there are separate, modality-dependent shape representations stored in the brain, or if metric information from multiple sources is integrated into one multi-sensory representation. It was concluded that there was a transfer of metric category knowledge across haptic and visual modalities, which indicated these forms of shape information are integrated into a shared multi-sensory representation. A related experiment testing whether haptic feedback can override visual information should be investigated with this in mind. Much like the application of the weighted model of cue combination, and the results of *Chapter 4*, this can be explored using a VR set-up with physical 3D prints of the virtual stimuli to test whether the weighting of visual and haptic cues are disproportionately combined. Participants would make size estimates for the stimulus in three conditions: (1) a similar set-up to *Chapter*

4, presenting participants with a virtual stimulus in VR, at differing distances in near space. (2) Void of visual cues, instead presented with a physical version of the stimulus for which the estimates are made. (3) Both the physical and the visual representations of the stimulus are presented for estimates to be based upon.

Based on the weighted cue-combination model, it would be assumed that the variability of size estimates in the combined visual and haptic condition would be the sum of the variance found in the estimates from the visual condition and haptic condition, combined. However, as per the results of Experiment 3 from *Chapter 4*, the two cues may be combined in a more optimal fashion. This could be observed in two ways: if the physical and virtual stimuli hold concordant dimensions, the variability in responses should diminish in the combined condition. Alternatively, if the dimensions of the two stimuli are inconsistent, any directional bias in the size estimates from the combined condition would indicate the weighting and prioritisation of either cue. For example if the physical stimulus was 20 cm and the visual was 10 cm, if size estimates surrounded 15 cm it can be deduced that the weighting of the two cues are equal, however if the estimates surround 17 cm this would suggest a bias towards the use of haptic cues rather than visual.

VR would be essential to this paradigm because, much like in *Chapter 3*, the ability to present stimuli at differing distances between each trial, while holding all extraneous variable constant, is effectively achieved through VR. With the inclusion of apparatus such as the Leap Motion Controller, which has been found to have minimal errors for tracking finger apertures (Hornsey & Hibbard, 2015), participants would be able to indicate their size estimates of the stimuli in this experiment. By using this modality of obtaining size estimates, it eliminates the chance of any unwanted visual size feedback from observing a virtual cross hair to indicate these settings.

Investigating the combination of auditory and visual cues would also be an area of research with high impact. A study by Zahorik (2002) used a psychophysical method to measure the cue-weighting process of intensity and direct-to-reverberant energy ratio in

a large, physical environment. It was found that participants consistently and exponentially under-estimated the true distance for the source of the sound; it was concluded that the cue-weighting process is flexible, in that it adapts to individual distance cues that vary, as a result of source properties and environmental conditions.

To further assess the impact sounds influence distance perception through the use of VR, an experiment in which participants make depth estimates of a well could be programmed. Participants should drop an item down a virtual well and estimate the depth to the bottom. In one condition the well would be lit so the dropped item can be observed to fall the entire depth; in another condition the well would not be lit, but a sound of the item hitting the bottom could be heard instead. Both the visible depth of the well and the time delay of the item hitting the ground would vary between trials, and the relationship between estimates and presented depth would be evaluated for both conditions. It would be expected that there would be an under-estimation of the distances in the visual condition (based on the findings of under-estimation of virtual distances presented throughout this thesis) as well as in the auditory condition (Ashmead et al., 1995; Zahorik, 2002; Zahorik et al., 2005).

The impact of these findings would be beneficial for understanding the contribution of auditory cues, and for simulations where sounds might be beneficial at aiding distance perception. For example, in instances where sonar is used to detect underwater masses using sound waves, and lidar is used with light waves, the way individuals weigh these cues when determining distances could have ramifications into the way submarine simulation training programmes are built.

7.2.4 Impact of Immersion on Perception

Slater et al. (1996) stated that immersion is an objective measure of the degree to which an environment is extensive, surrounding, inclusive, and vivid. In this study they aimed to investigate the impact of immersion by having participants complete a task with an egocentric viewpoint (using a 3D stereoscopic helmet to elicit a strong sense of

immersion) and an exocentric viewpoint (2D TV comparison condition). It was found that immersion improved performance in their task, perhaps due to additional visual information available from the type of display.

In another domain, VR exposure therapy (VRET) has been used to reduce the negative affective symptoms experienced by phobia patients, in a controlled, virtual environment. A study by Krijn et al. (2004) used VRET to treat participants diagnosed with acrophobia. Three virtual scenarios were created and were each observed for one-hour by patients; questionnaires and a behavioural avoidance test were administered after each session, in addition to a follow-up six months later to assess any lasting effects of the treatment. Comparison of the post-test and follow-up revealed a desirable time effect, indicating that the improvements from the VRET remained stable six months later. Additionally, VRET results were significantly different from a control group on all levels. Although the results of this experiment were favourable, there is always room for improvement, and the longevity of the therapy may be improved further by using more realistic and immersive software. The finding of this study, where VRET was successful in reducing symptoms of phobias, have been found for numerous other forms of anxiety disorders: for instance, Parsons and Rizzo (2008) conducted a meta-analysis which analysed published papers on participants who had been diagnosed with a specific anxiety disorder. Significant results were found across the range of different disorders tested, suggesting this is a worthwhile application for VR.

Another area in which immersion is likely to have an effect is in the ability to take another person's perspective in spatial cognitions. An experiment to test the way immersion is influenced by presentation medium and sense of involvement would be beneficial, especially for therapy-related applications of virtual environments. In this study, three conditions invoking different levels of involvement would be tested on a perspective-taking task; it is predicted that the ability to take on another perspective would be increased by an enhanced sense of engagement. The sense of involvement should be determined via a questionnaire completed after each of the three conditions: (1) en-

vironment presented on a high definition 2D computer screen using a chin rest; (2) presented in a VR headset using a chin rest; (3) presented in VR along with visible hands matching the controller positions, background audio, and free head movement from the observer. The collected metrics would be reaction time, response (correct, incorrect), and overall involvement rating for each condition. The predicted relationships between the subjective and objective results would be that the highest sense of involvement is to be experienced in the final condition, which would also provide the lowest reaction times for correct responses. Therapy where taking another individual's perspective has been found to be beneficial for future-self continuity (Ganschow et al., 2021), cognitive reflection (Kellogg, 2014), and help facilitate insights into challenging clinical encounters (Safran et al., 2008).

7.3 Conclusions

In cognitive sciences the use of specialist equipment has been grandfathered; offering high reliability at a large financial cost and specialist training for use and programming set-ups. However this thesis has presented an alternative, providing evidence of multiple contexts for which VR can be implemented into scientific research and evidenced how it can be used to overcome previously restricting obstacles.

The experimental chapters have found evidence for reliable presentation of virtual environments, with accuracy of displayed distances comparable with the aforementioned traditional, carefully calibrated equipment. Although distances are generally underestimated, these studies have shown that by adding visual cues and lighting techniques, the mis-estimation can be reduced. This thesis has discussed the benefits of using such equipment, and provided a guide for setting up a psychophysical VR experiment.

Evidence for the acceptance of VR equipment in daily life has been explored here; by gathering user experiences and using qualitative methodology, the constraints of the devices and improvements have been collated in order to understand the reasoning behind why some individuals are resistant to use of the equipment. The appropriateness

of using VR technology for future research is widely acknowledged as being important, which reinforces the notion that this equipment is suitable and will be accepted by majority of the public.

Appendix A

Chapter 6 Questionnaire

1. `Consent` - options to `Accept` or `Decline`; if `Decline` selected skip to end of questionnaire

---

2. `What is your age (in years)?` - free text, numerical
3. `Which generation do you associate yourself with?` - options `Generation Z, Millennial, Generation X, Baby boomer, Silent generation, I don't know`
4. `What is your gender?` options `Male, Female, Other`
5. `Do you have 20/20 vision, or better?` - options `Yes, No, Don't know because I've never had an eye test`
   (a) display if `Q5 = No` - `Do you wear glasses?` - options `Yes, No`
   (b) display if `Q5 = No` - `Do you wear contact lenses?` - options `Yes, No`

---

6. `Do you own a VR device?` - options `Yes, No`
   (a) display if `Q8 = Yes` - `Which VR headset do you own?` - free text
7. `Have you tried VR before?` - options `Yes, No`
   (a) display if `Q7 = No` - `For what reason(s) have you not tried VR before?` - free text
   (b) display if `Q7 = Yes` - `How often do you use VR?` - options `Rarely, Sometimes, Often`
   (c) display if `Q7 = Yes` - `Do you enjoy using VR?` - options `Yes, No`

(d) display if Q7 = Yes - What purposes do you use VR for? (select all that apply) - options Educational, Recreational, Business, Other: (free text)

(e) display if Q7 = Yes - When was the last time you used VR? - options Today, Yesterday, A few days ago, A week ago, A few weeks ago, Over a month ago, So long go I can't remember when it was

(f) display if Q7 = Yes - Have you ever had the distance between your eyes/pupils measured? - options Yes, No

i. display if Q7f = Yes - Do you adjust the setting on the VR device to match the distance measured between your eyes/pupils? - options Yes, No

(g) display if Q7 = Yes and Q5a = Yes - Do you wear your glasses when using VR? - options Yes, No

8. What do you associate VR with? - free text

9. What are your thoughts on the VR headsets themselves? - items Lightweight, Easy to use, Bulky, Useful, Complicated to set up, Unnecessary, Enjoyable to use, I will own one in the next few years; options Strongly agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Strongly disagree

10. Rank these sensory factors for what is most important (1) to least important (4) in VR for you: - options Accurate head-tracking, Visual graphics, High-quality sound, Controller/haptic feedback

11. What is the most important to you when it comes to a VR headset? - options Price, Weight, Available content, Popularity, Customisability

---

12. Have you ever had a text to check if you were able to perceive 3D stimuli? - options Yes, No

13. Do you experience travel-induced motion sickness? - options Yes, No

14. Do you experience technology-induced motion sickness? - options Yes, No

15. Do you experience any visual impairment (not including wearing glasses/contact lenses)? - options Yes, No

16. Do you experience any auditory impairment? - options Yes, No

---

17. Mobile phones can record data through cameras, microphones, location services, and browser history. Some of these features can be turned on and off, however many downloadable applications require access to these features in order to work. Do you have any opinions on the data that is collected on you, from your mobile phone? - free text

18. Facial recognition is the technology used in mobile phones which allows for the phone to recognise when the owner is using the phone. This feature can be used for security purposes (such as automatically unlocking the phone when the owner is looking at the screen), but can also be used for entertainment purposes (creating personalised avatars with filters). What are your thoughts on facial recognition on mobile devices? - options Extremely appropriate, Somewhat appropriate, Neither appropriate nor inappropriate, Somewhat inappropriate, Extremely inappropriate

19. What are your thoughts on facial recognition on VR devices? - options Extremely appropriate, Somewhat appropriate, Neither appropriate nor inappropriate, Somewhat inappropriate, Extremely inappropriate

20. Eye-tracking is the technology used in research experiments to know where the individual is looking within a scene that is presented to them. This can be used to see where a person's focus is, and for how long specific areas hold their attention. What are your thoughts on eye-tracking during research experiments? - options Extremely appropriate, Somewhat appropriate, Neither appropriate nor inappropriate, Somewhat inappropriate, Extremely inappropriate

21. What are your thoughts on eye-tracking within VR devices? - options Extremely appropriate, Somewhat appropriate, Neither appropriate nor inappropriate, Somewhat inappropriate, Extremely inappropriate

22. Content personalisation is a marketing strategy that presents content which is deemed relevant for a viewer, based on algorithms using previous interests and search history. What are your thoughts on content personalisation within social media advertisements? - options Extremely appropriate, Somewhat appropriate,

Neither appropriate nor inappropriate, Somewhat inappropriate, Extremely inappropriate

23. What are your thoughts on content personalisation within VR content? - options Extremely appropriate, Somewhat appropriate, Neither appropriate nor inappropriate, Somewhat inappropriate, Extremely inappropriate
24. Do you have any opinions on the data that is collected on you, from a VR headset? - free text
25. Can you see VR becoming a daily commodity in the next few years? - options Yes, No
26. What are your thoughts on VR becoming a daily commodity? - free text
27. What ways do you think VR could benefit you specifically? - free text
28. What precautions do you think need to be taken when using VR? - free text
29. What precautions do you think need to be taken when creating a new VR headset? - free text
30. What improvements do you think need to be made to the current VR headsets? - free text
31. Rank the brands for who has created the best VR headset: - options Sony, HTC, Oculus, Samsung, Valve
32. What is your answer to the previous question based on? - options Own experience, Knowledge on them Generally, Other: free text

---

33. Have you ever used augmented reality? - options Yes, No
34. Have you ever used mixed reality? - options Yes, No
35. Do you enjoy theatre productions? - options Yes, No
36. Do you enjoy watching movies or TV programmes? - options Yes, No
37. Do you enjoy online gaming? - options Yes, No
38. Do you enjoy offline gaming? - options Yes, No

39. `Do you enjoy board games?` - options `Yes, No`

40. `Which gaming consoles do play on?` - options `PC, Microsoft Xbox, Sony PlayStation, Mobile phone/tablet, Nintendo, None, Other:` free text

Appendix B

Databases for Raw Data

B.1 Size and Shape Constancy in Consumer Virtual Reality

Available at: http://researchdata.essex.ac.uk/87/

B.2 Contribution of Differing Visual Cues on Distance Perception in Virtual Reality

Available at: http://researchdata.essex.ac.uk/128/

B.3 The Role of Cast Shadows in the Perception of Distance

Available at: http://researchdata.essex.ac.uk/143/

B.4 Exploration of Perceptions Surrounding Virtual Reality Technology

Available at: http://researchdata.essex.ac.uk/142/

Bibliography

Abrash, M. (2014). What vr could, should, and almost certainly will be within two years. Retrieved October 14, 2021, from https://www.youtube.com/watch?v=G-2dQoeqVVo%5C&t=3s

Adamovich, S. V., Fluet, G. G., Tunik, E. & Merians, A. S. (2009). Sensorimotor training in virtual reality: A review. *NeuroRehabilitation*, *25*(1), 29–44.

Adams, W. J., Graf, E. W. & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature neuroscience*, *7*(10), 1057–1058.

Adorama. (2021). Types of lighting in film: Basic techniques to know. Retrieved May 28, 2021, from https://www.adorama.com/alc/basic-cinematography-lighting-techniques/

Agarwal, A., Singh, R. & Toshniwal, D. (2018). Geospatial sentiment analysis using twitter data for uk-eu referendum. *Journal of Information and Optimization Sciences*, *39*(1), 303–317.

Agarwal, R. & Prasad, J. (1997). The role of innovation characteristics and perceived voluntariness in the acceptance of information technologies. *Decision sciences*, *28*(3), 557–582.

Aghasi, A., Heshmat, B., Wei, L. & Tian, M. (2021). Optimal allocation of quantized human eye depth perception for multi-focal 3d display design. *Optics Express*, *29*(7), 9878–9896.

Agrawal, S., Simon, A., Bech, S., Bæntsen, K. & Forchhammer, S. (2020). Defining immersion: Literature review and implications for research on audiovisual experiences. *Journal of the Audio Engineering Society*, *68*(6), 404–417.

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *44*(2), 139–160.

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes*, *50*(2), 179–211.

Akeley, K., Watt, S. J., Girshick, A. R. & Banks, M. S. (2004). A stereo display prototype with multiple focal distances. *ACM transactions on graphics (TOG)*, *23*(3), 804–813.

Alais, D. & Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Current biology*, *14*(3), 257–262.

Allison, R. S., Gillam, B. J. & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*, *9*(1), 10–10.

Alvarado, R. G., Holzapfel, R. H. & Parra, J. C. (1999). Virtual design for innovative timber structures.

Amoako-Gyampah, K. (2007). Perceived usefulness, user involvement and behavioral intention: An empirical study of erp implementation. *Computers in human behavior*, *23*(3), 1232–1248.

Angarita, A., Hernandez, A., Carmichael, C., Uribe-Quevedo, A., Rueda, C. & Salinas, S. A. (2019). Increasing virtual reality immersion through smartwatch lower limb motion tracking, In *International conference on human-computer interaction*. Springer.

Arnold, D. H., Petrie, K., Murray, C. & Johnston, A. (2019). Suboptimal human multisensory cue combination. *Scientific Reports*, *9*(1), 1–11.

Arthur, K. W. (2000). *Effects of field of view on performance with head-mounted displays*. The University of North Carolina at Chapel Hill.

Asch, S. E. (1956). Studies of independence and conformity: I. a minority of one against a unanimous majority. *Psychological monographs: General and applied*, *70*(9), 1.

Asher, J. M. & Hibbard, P. B. (2018). First-and second-order contributions to depth perception in anti-correlated random dot stereograms. *Scientific reports*, *8*(1), 1–19.

Asher, J. M., Romei, V. & Hibbard, P. B. (2019). Spatial frequency tuning and transfer of perceptual learning for motion coherence reflects the tuning properties of global motion processing. *Vision*, *3*(3), 44.

Ashmead, D. H., Davis, D. L. & Northington, A. (1995). Contribution of listeners' approaching motion to auditory distance perception. *Journal of experimental psychology: Human perception and performance*, *21*(2), 239.

Ayata, D., Yaslan, Y. & Kamaşak, M. (2017). Emotion recognition via galvanic skin response: Comparison of machine learning algorithms and feature extraction methods. *Istanbul University-Journal of Electrical & Electronics Engineering*, *17*(1), 3147–3156.

Backus, B. T., Banks, M. S., Van Ee, R. & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision research*, *39*(6), 1143–1170.

Baltrusaitis, J. (2021). Number of gdpr fines surge by 113% in a year despite strict regulations (Finbold, Ed.). Retrieved October 21, 2021, from https://finbold.com/number-of-gdpr-fines-2021/

Banks, M. S., Hoffman, D. M., Kim, J. & Wetzstein, G. (2016). 3d displays.

Banks, M. S., Read, J. C., Allison, R. S. & Watt, S. J. (2012). Stereoscopy and the human visual system. *SMPTE motion imaging journal*, *121*(4), 24–43.

Barfield, W. & Rosenberg, C. (1995). Judgments of azimuth and elevation as a function of monoscopic and binocular depth cues using a perspective display. *Human Factors*, *37*(1), 173–181.

Battaglia, P. W., Jacobs, R. A. & Aslin, R. N. (2003). Bayesian integration of visual and auditory signals for spatial localization. *Josa a*, *20*(7), 1391–1397.

Berkman, M. & Akan, E. (2019). Presence and immersion in virtual reality (N. Lee, Ed.). In N. Lee (Ed.), *Encyclopedia of computer graphics and games*. Cham, Springer International Publishing. https://doi.org/10.1007/978-3-319-08234-9_162-1

Berman, R., Piller, M. & Taylor, J. (1995). Star trek: Voyager.

Bernstein, H. (1971). Modernization theory and the sociological study of development. *The Journal of Development Studies*, *7*(2), 141–160.

Bertrand, M. & Bouchard, S. (2008). Applying the technology acceptance model to vr with people who are favorable to its use. *Journal of Cyber Therapy & Rehabilitation*, *1*(2), 200–210.

Beusmans, J. M. (1998). Perceived object shape affects the perceived direction of self-movement. *Perception*, *27*(9), 1079–1085.

Bhalla, M. & Proffitt, D. R. (1999). Visual–motor recalibration in geographical slant perception. *Journal of experimental psychology: Human perception and performance*, *25*(4), 1076.

Bichlmeier, C. & Navab, N. (2006). Virtual window for improved depth perception in medical ar, In *International workshop on augmented reality environments for medical imaging and computer-aided surgery (ami-arcs)*. Citeseer.

Bideau, B., Kulpa, R., Vignais, N., Brault, S., Multon, F. & Craig, C. (2009). Using virtual reality to analyze sports performance. *IEEE Computer Graphics and Applications*, *30*(2), 14–21.

Bingham, G. P., Bradley, A., Bailey, M. & Vinner, R. (2001). Accommodation, occlusion, and disparity matching are used to guide reaching: A comparison of actual versus virtual environments. *Journal of experimental psychology: human perception and performance*, *27*(6), 1314.

Bishop, P. (1989). Vertical disparity, egocentric distance and stereoscopic depth constancy: A new interpretation. *Proceedings of the Royal Society of London B, 237*, *445469*.

Bjørner, T., Magnusson, A. & Nielsen, R. (2016). How to describe and measure obstacles of narrative immersion in a film? *Nordicom Review*, *37*(1), 1–17. https://content.sciendo.com/view/journals/nor/37/1/article-p1.xml

Blei, D. & Lafferty, J. (2006). Correlated topic models. *Advances in neural information processing systems*, *18*, 147.

Bodenheimer, B., Meng, J., Wu, H., Narasimham, G., Rump, B., McNamara, T. P., Carr, T. H. & Rieser, J. J. (2007). Distance estimation in virtual and real environments using bisection, In *Proceedings of the 4th symposium on applied perception in graphics and visualization*. ACM.

Bodyswaps. (2021). Changing behaviours - for good. Retrieved September 3, 2021, from https://bodyswaps.co/

Bradshaw, M. F., Parton, A. D. & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision research*, *40*(27), 3725–3734.

Brasel, S. A. & Gips, J. (2014). Tablets, touchscreens, and touchpads: How varying touch interfaces trigger psychological ownership and endowment. *Journal of Consumer Psychology*, *24*(2), 226–233.

Brenner, E. & Van Damme, W. J. (1998). Judging distance from ocular convergence. *Vision research*, *38*(4), 493–498.

Brenner, E. & van Damme, W. J. (1999). Perceived distance, shape and size. *Vision research*, *39*(5), 975–986.

Brown, E. & Cairns, P. (2004). A grounded investigation of game immersion, In *Chi'04 extended abstracts on human factors in computing systems*.

Brown, R. (2021). Forte vfx1 vs oculus rift (comparison). Retrieved July 20, 2021, from https://vr-compare.com/compare?h1=bpzF1luyF%5C&h2=-CBHn3rsnf_

Bruner II, G. C. & Kumar, A. (2005). Explaining consumer acceptance of handheld internet devices. *Journal of business research*, *58*(5), 553–558.

Bruner, J. S. & Goodman, C. C. (1947). Value and need as organizing factors in perception. *The journal of abnormal and social psychology*, *42*(1), 33.

Brunstrom, J. M. (2011). The control of meal size in human subjects: A role for expected satiety, expected satiation and premeal planning. *Proceedings of the Nutrition Society*, *70*(2), 155–161.

Bülthoff, H. H. & Mallot, H. A. (1988). Integration of depth modules: Stereo and shading. *Josa a*, *5*(10), 1749–1758.

Burns, E., Razzaque, S., Panter, A. T., Whitton, M. C., McCallus, M. R. & Brooks, F. P. (2005). The hand is slower than the eye: A quantitative exploration of visual dominance over proprioception, In *Ieee proceedings. vr 2005. virtual reality, 2005.* IEEE.

Carlson, V. (1960). Overestimation in size-constancy judgments. *The American journal of psychology*, *73*(2), 199–213.

Casati, R. (2004). The shadow knows: A primer on the informational structure of cast shadows. *Perception*, *33*(11), 1385–1396.

Chalmers, E. L. (1952). Monocular and binocular cues in the perception of size and distance. *The American journal of psychology*, *65*(3), 415–423.

Chau, P. Y. & Lai, V. S. (2003). An empirical investigation of the determinants of user acceptance of internet banking. *Journal of organizational computing and electronic commerce*, *13*(2), 123–145.

Chirico, A., Lucidi, F., De Laurentiis, M., Milanese, C., Napoli, A. & Giordano, A. (2016). Virtual reality in health system: Beyond entertainment. a mini-review on the efficacy of vr during cancer treatment. *Journal of cellular physiology*, *231*(2), 275–287.

Choi, J. K. & Ji, Y. G. (2015). Investigating the importance of trust on adopting an autonomous vehicle. *International Journal of Human-Computer Interaction*, *31*(10), 692–702.

Chopin, A., Bavelier, D. & Levi, D. M. (2019). The prevalence and diagnosis of "stereo-blindness": A best evidence synthesis. *Journal of Vision*, *19*(10), 262b–262b.

Chung, J. E., Park, N., Wang, H., Fulk, J. & McLaughlin, M. (2010). Age differences in perceptions of online community participation among non-users: An extension of the technology acceptance model. *Computers in Human Behavior*, *26*(6), 1674–1684.

Cochran, W. G. (1937). Problems arising in the analysis of a series of similar experiments. *Supplement to the Journal of the Royal Statistical Society*, *4*(1), 102–118.

Colby, S. & Ortman, J. M. (2014). *The baby boom cohort in the united states: 2012 to 2060*. US Department of Commerce, Economics; Statistics Administration, US ...

Collewijn, H., Van der Steen, J. & Van Rijn, L. (1991). Binocular eye movements and depth perception. *Representations of vision, trends and tacit assumptions in vision research*, 165–183.

Collier, E. S. (2018). *The illusion of action-specific scaling effects: Action capacity does not directly influence spatial perception*. The University of Liverpool (United Kingdom).

Corujeira, J. G. & Oakley, I. (2013). Stereoscopic egocentric distance perception: The impact of eye height and display devices, In *Proceedings of the acm symposium on applied perception*.

Creem-Regehr, S. H. & Kunz, B. R. (2010). Perception and action. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 800–810.

Creem-Regehr, S. H., Stefanucci, J. K. & Thompson, W. B. (2015). Perceiving absolute scale in virtual environments: How theory and application have mutually informed

the role of body-based perception, In *Psychology of learning and motivation*. Elsevier.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A. & Thompson, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual indoor environments. *Perception*, *34*(2), 191–204.

Crompton, A. & Brown, F. (2006). Distance estimation in a small-scale environment. *Environment and Behavior*, *38*(5), 656–666.

Cronenberg, D. (1999). Existenz.

Crutchfield, P., Pazdernik, V., Hansen, G., Malone, J. & Wagenknecht, M. (2018). Being hungry affects oral size perception. *i-Perception*, *9*(3), 2041669518777513.

Cutting, J. E. & Vishton, P. M. (1995). Perceiving layout and knowing distances: The integration, relative potency, and contextual use of different information about depth, In *Perception of space and motion*. Elsevier.

Czaja, S. J., Charness, N., Fisk, A. D., Hertzog, C., Nair, S. N., Rogers, W. A. & Sharit, J. (2006). Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement (create). *Psychology and aging*, *21*(2), 333.

Dahal, B., Kumar, S. A. & Li, Z. (2019). Topic modeling and sentiment analysis of global climate change tweets. *Social Network Analysis and Mining*, *9*(1), 1–20.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319–340.

Davis, F. D. Et al. (1989a). Technology acceptance model: Tam. *Al-Suqri, MN, Al-Aufi, AS: Information Seeking Behavior and Technology Adoption, S*, 205–219.

Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989b). User acceptance of computer technology: A comparison of two theoretical models. *Management science*, *35*(8), 982–1003.

Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1992). Extrinsic and intrinsic motivation to use computers in the workplace 1. *Journal of applied social psychology*, *22*(14), 1111–1132.

De Paolis, L. T. & De Luca, V. (2019). Augmented visualization with depth perception cues to improve the surgeon's performance in minimally invasive surgery. *Medical & biological engineering & computing*, *57*(5), 995–1013.

Dee, H. M. & Santos, P. E. (2011). The perception and content of cast shadows: An interdisciplinary review. *Spatial Cognition & Computation*, *11*(3), 226–253.

Defreitas, M. (2018). SCAD-stranauts: SCAD day's destination imagination. Retrieved May 28, 2020, from https://scaddistrict.com/2018/04/19/scad-stranauts-scad-days-destination-imagination/

Dekogon Studios. (2018). *Construction site vol. 2 - tools, parts, and machine props* [Unreal Engine Marketplace]. Retrieved October 22, 2020, from https://marketplace-website-node-launcher-prod.ol.epicgames.com/ue/marketplace/en-US/product/construction-site-vol-2-tools-parts-and-machine-props

Deng, Z. & Interrante, V. (2019). Am i floating or not?: Sensitivity to eye height manipulations in hmd-based immersive virtual environments, In *Acm symposium on applied perception 2019*.

Dimenco. (2022). Introducing the sr pro display. Retrieved May 22, 2022, from https://www.dimenco.eu/sr-pro-display

Dixon, M. W., Wraga, M., Proffitt, D. R. & Williams, G. C. (2000). Eye height scaling of absolute size in immersive and nonimmersive displays. *Journal of Experimental Psychology: Human Perception and Performance*, *26*(2), 582.

Dodgson, N. A. (2004). Variation and extrema of human interpupillary distance, In *Stereoscopic displays and virtual reality systems xi*. International Society for Optics and Photonics.

Douglas, Y. & Hargadon, A. (2000). The pleasure principle: Immersion, engagement, flow, In *Proceedings of the eleventh acm on hypertext and hypermedia*.

Drascic, D. (1991). Skill acquisition and task performance in teleoperation using monoscopic and stereoscopic video remote viewing, In *Proceedings of the human factors society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Drascic, D. & Milgram, P. (1996). Perceptual issues in augmented reality, In *Stereoscopic displays and virtual reality systems iii*. International Society for Optics and Photonics.

Du, H. & Li, J. (2021). Application of computer virtual reality technology in university volleyball teaching, In *International conference on application of intelligent systems in multi-modal information analytics*. Springer.

Duane, A. (1922). Studies in monocular and binocular accommodation, with their clinical application. *Transactions of the American Ophthalmological Society*, *20*, 132.

Dyde, R., MacKenzie, K. & Harris, L. (2011). How well do you know the back of your hand? reaction time to identify a rotated hand silhouette depends on whether it is interpreted as a palm or back view. *Journal of Vision*, *11*(11), 868–868.

Eccleston, C. & Crombez, G. (1999). Pain demands attention: A cognitive–affective model of the interruptive function of pain. *Psychological bulletin*, *125*(3), 356.

Eggleston, R. G., Janson, W. P. & Aldrich, K. A. (1996). Virtual reality system effects on size-distance judgements in a virtual environment, In *Proceedings of the ieee 1996 virtual reality annual international symposium*. IEEE.

Elias, Z. M., Batumalai, U. M. & Azmi, A. N. H. (2019). Virtual reality games on accommodation and convergence. *Applied ergonomics*, *81*, 102879.

Ellis, R. D. & Allaire, J. C. (1999). Modeling computer interest in older adults: The role of age, education, computer knowledge, and computer anxiety. *Human Factors*, *41*(3), 345–355.

Ellis, S. R. & Menges, B. M. (1998). Localization of virtual objects in the near visual field. *Human factors*, *40*(3), 415–431.

Elshafey, A., Saar, C. C., Aminudin, E. B., Gheisari, M. & Usmani, A. (2020). Technology acceptance model for augmented reality and building information modeling integration in the construction industry. *J. Inf. Technol. Constr.*, *25*, 161–172.

Emperore, K. & Sherry, D. (2015). *Unreal engine physics essentials*. Packt Publishing Ltd.

Engonopoulos, N., Lazaridou, A., Paliouras, G. & Chandrinos, K. (2011). Els: A word-level method for entity-level sentiment analysis, In *Proceedings of the international conference on web intelligence, mining and semantics*.

Epic Games. (2021). Unreal engine 5 early access. Retrieved November 3, 2021, from https://www.unrealengine.com/en-US/unreal-engine-5

Ernst, M. O. & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433.

European Parliament And Council Of The European Union. (2016). Regulation (eu) 2016/679 of the european parliament and of the council. Retrieved October 15, 2021, from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

FailArmy. (2021). Game over — funny virtual reality fails. Retrieved October 21, 2021, from https://www.youtube.com/watch?v=o6wO0pL0tQs

Farra, N., Challita, E., Abou Assi, R. & Hajj, H. (2010). Sentence-level and document-level sentiment mining for arabic texts, In *2010 ieee international conference on data mining workshops*. IEEE.

Fastl, H. & Zwicker, E. (2006). *Psychoacoustics: Facts and models* (Vol. 22). Springer Science & Business Media.

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82–89.

Fellows, I. (2018). Package 'wordcloud' [CRAN]. Retrieved October 12, 2021, from https://cran.r-project.org/web/packages/wordcloud/wordcloud.pdf

Fernandes, A. S. & Feiner, S. K. (2016). Combating vr sickness through subtle dynamic field-of-view modification, In *2016 ieee symposium on 3d user interfaces (3dui)*. IEEE.

Fetsch, C. R., Pouget, A., DeAngelis, G. C. & Angelaki, D. E. (2012). Neural correlates of reliability-based cue weighting during multisensory integration. *Nature neuroscience*, *15*(1), 146–154.

Fishbein, M. & Ajzen, I. (1977). Belief, attitude, intention, and behavior: An introduction to theory and research. *Philosophy and Rhetoric*, *10*(2).

FLAIM. (2021). *Flaim trainer^{TM}, the world's first immersive technology enabled firefighter training solution*. Retrieved September 3, 2021, from https://flaimsystems.com/products/trainer

Foley, J. M. (1980). Binocular distance perception. *Psychological review*, *87*(5), 411.

Foley, J. M. (1991). Binocular space perception. *Binocular vision*, *9*, 75–92.

Freeman, D. (2008). Studying and treating schizophrenia using virtual reality: A new paradigm. *Schizophrenia bulletin*, *34*(4), 605–610.

Freina, L. & Ott, M. (2015). A literature review on immersive virtual reality in education: State of the art and perspectives, In *The international scientific conference elearning and software for education*. " Carol I" National Defence University.

Frisby, J. P., Buckley, D. & Horsman, J. M. (1995). Integration of stereo, texture, and outline cues during pinhole viewing of real ridge-shaped objects and stereograms of ridges. *Perception*, *24*(2), 181–198.

Game. (2016). Htc vive full trailer. https://www.youtube.com/watch?v=i1r76omNeI8

Ganschow, B., Cornet, L., Zebel, S. & Van Gelder, J.-L. (2021). Looking back from the future: Perspective taking in virtual reality increases future self-continuity. *Frontiers in psychology*, *12*.

Gantman, A. P. & Van Bavel, J. J. (2014). The moral pop-out effect: Enhanced perceptual awareness of morally relevant stimuli. *Cognition*, *132*(1), 22–29.

Gao, Y., Peillard, E., Normand, J.-M., Moreau, G., Liu, Y. & Wang, Y. (2020). Influence of virtual objects' shadows and lighting coherence on distance perception in optical see-through augmented reality. *Journal of the Society for Information Display*, *28*(2), 117–135.

Gartner, W. C. (1994). Image formation process. *Journal of travel & tourism marketing*, *2*(2-3), 191–216.

Gearbox Software. (2019). *Borderlands 3* [2K]. Retrieved May 28, 2020, from https://www.epicgames.com/store/en-US/product/borderlands-3/home

Georgeson, M. A., Yates, T. A. & Schofield, A. J. (2009). Depth propagation and surface construction in 3-d vision. *Vision research*, *49*(1), 84–95.

Geuss, M. N., Stefanucci, J. K., Creem-Regehr, S. H., Thompson, W. B. & Mohler, B. J. (2015). Effect of display technology on perceived scale of space. *Human factors*, *57*(7), 1235–1247.

Ghisletta, P., Renaud, O., Fagot, D., Lecerf, T. & De Ribaupierre, A. (2018). Age and sex differences in intra-individual variability in a simple reaction time task. *International Journal of Behavioral Development*, *42*(2), 294–299.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton, Mifflin; Company.

Gilson, S. J., Fitzgibbon, A. W. & Glennerster, A. (2008). Spatial calibration of an optical see-through head-mounted display. *Journal of neuroscience methods*, *173*(1), 140–146.

Gilson, S. J., Fitzgibbon, A. W. & Glennerster, A. (2011). An automated calibration method for non-see-through head mounted displays. *Journal of neuroscience methods*, *199*(2), 328–335.

Girshick, A., Akeley, K., Watt, S. & Banks, M. (2004). Improved stereoscopic performance with consistent vergence and accommodative cues in a novel 3-d display.

Glennerster, A., Tcheang, L., Gilson, S. J., Fitzgibbon, A. W. & Parker, A. J. (2006). Humans ignore motion and stereo cues in favor of a fictional stable world. *Current Biology*, *16*(4), 428–432.

Glennerster, A., Rogers, B. J. & Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. *Vision research*, *36*(21), 3441–3456.

Gogel, W. C. (1965). Equidistance tendency and its consequences. *Psychological Bulletin*, *64*(3), 153.

Gogel, W. C. (1977). An indirect measure of perceived distance from oculomotor cues. *Perception & Psychophysics*, *21*(1), 3–11.

Gong, H., Watanabe, N. M., Soebbing, B. P., Brown, M. T. & Nagel, M. S. (2021). Do consumer perceptions of tanking impact attendance at national basketball association games? a sentiment analysis approach. *Journal of Sport Management*, *35*(3), 254–265.

Google. (2021). Google opinion rewards. Retrieved October 18, 2021, from https://surveys.google.com/google-opinion-rewards/

Goutcher, R. & Hibbard, P. B. (2014). Mechanisms for similarity matching in disparity measurement. *Frontiers in Psychology*, *4*, 1014.

Graham, R. B., Agnew, M. J. & Stevenson, J. M. (2009). Effectiveness of an on-body lifting aid at reducing low back physical demands during an automotive assembly task: Assessment of emg response and user acceptability. *Applied Ergonomics*, *40*(5), 936–942.

Groen, J. & Werkhoven, P. J. (1998). Visuomotor adaptation to virtual hand position in interactive virtual environments. *Presence*, *7*(5), 429–446.

Grubert, J., Witzani, L., Ofek, E., Pahud, M., Kranz, M. & Kristensson, P. O. (2018). Effects of hand representations for typing in virtual reality, In *2018 ieee conference on virtual reality and 3d user interfaces (vr)*. IEEE.

Grün, B. & Hornik, K. (2011). topicmodels: An R package for fitting topic models. *Journal of Statistical Software*, *40*(13), 1–30.

Ha, S. & Stoel, L. (2009). Consumer e-shopping acceptance: Antecedents in a technology acceptance model. *Journal of business research*, *62*(5), 565–571.

Haines, A. E., Hornsey, R. L. & Hibbard, P. B. (2015). Quality, quantity and precision of depth perception in stereoscopic displays, In *2015 international conference on 3d imaging (ic3d)*. IEEE.

Hamamura, T. (2012). Are cultures becoming individualistic? a cross-temporal comparison of individualism–collectivism in the united states and japan. *Personality and social psychology review*, *16*(1), 3–24.

Hanley, J. R. & Bourgaize, J. (2018). Similarities between the irrelevant sound effect and the suffix effect. *Memory & cognition*, *46*(6), 841–848.

HaptiX. (2021). Hapticx gloves dk2. Retrieved November 11, 2021, from https://haptx.com/

Hartle, B. & Wilcox, L. M. (2021). Cue vetoing in depth estimation: Physical and virtual stimuli. *Vision Research*, *188*, 51–64.

Hartle, B. & Wilcox, L. M. (2022). Stereoscopic depth constancy for physical objects and their virtual counterparts. *Journal of Vision*, *22*(4), 9–9.

Hawaii News Now. (2020). No shadows: Parts of hawaii to experience 'lahaina noon'. Retrieved December 9, 2021, from https://www.hawaiinewsnow.com/2020/05/14/no-shadows-parts-hawaii-experience-lahaina-noon/

Hawkins, D. M. (1981). A cusum for a scale parameter. *Journal of Quality Technology*, *13*(4), 228–231.

Hawkins, D. M. & Wixley, R. (1986). A note on the transformation of chi-squared variables to normality. *The American Statistician*, *40*(4), 296–298.

Healey, N. (2016). *The ghost of vr past: The long history of tech's latest love affair (pictures)*. Retrieved March 12, 2021, from https://www.cnet.com/pictures/the-ghost-of-vr-past-the-long-history-of-techs-latest-darling/

Held, R. & Durlach, N. (1989). Telepresence, time delay, and adaptation. *Pictorial communication in virtual and real environments*.

Held, R. T. & Banks, M. S. (2008). Misperceptions in stereoscopic displays: A vision science perspective, In *Proceedings of the 5th symposium on applied perception in graphics and visualization*. ACM.

Held, R. T., Cooper, E. A. & Banks, M. S. (2012). Blur and disparity are complementary cues to depth. *Current biology*, *22*(5), 426–431.

Held, R. T., Cooper, E. A., O'brien, J. F. & Banks, M. S. (2010). Using blur to affect perceived distance and size. *ACM transactions on graphics*, *29*(2).

Henriksson, E. A. (2018). Data protection challenges for virtual reality applications. *Interactive Entertainment Law Review*, *1*(1), 57–61.

Hernandez, B., Jimenez, J. & Martín, M. J. (2009). Adoption vs acceptance of e-commerce: Two different decisions. *European Journal of Marketing*.

Hibbard, P. (2022 in press). Virtual reality for vision sciences. *Current Trends in Behavioral Neuroscience*, *0*(0), 0.

Hibbard, P. B. (in press). Estimating the contributions of pictorial, motion and binocular cues to the perception of distance.

Hibbard, P. B. (2008a). Binocular energy responses to natural images. *Vision Research*, *48*(12), 1427–1439.

Hibbard, P. B. (2008b). Can appearance be so deceptive? representationalism and binocular vision. *Spatial Vision*.

Hibbard, P. B., Goutcher, R., O'Kane, L. M. & Scarfe, P. (2012). Misperception of aspect ratio in binocularly viewed surfaces. *Vision research*, *70*, 34–43.

Hibbard, P. B., Haines, A. E. & Hornsey, R. L. (2017a). Magnitude, precision, and realism of depth perception in stereoscopic vision. *Cognitive Research: Principles and Implications*, *2*(1), 1–11.

Hibbard, P. B., van Dam, L. C. & Scarfe, P. (2020). The implications of interpupillary distance variability for virtual reality, In *2020 international conference on 3d immersion (ic3d)*. IEEE.

Hibbard, P., Goutcher, R., Khan, N. & Hornsey, R. (2017b). Manipulations of local, but not global, luminance gradients affect judgements of depth magnitude. *Journal of Vision*, *17*(10), 1045–1045.

Hibbard, P., Hornsey, R. L. & Khan, N. Z. (2017c). The contribution of binocular and motion cues to depth quality in complex naturalistic scenes, In *Perception*. SAGE Publications LTD 1 Olivers Yard, 55 City Road, London EC1Y 1SP, England.

Higashiyama, A. (1996). Horizontal and vertical distance perception: The discorded-orientation theory. *Perception & Psychophysics*, *58*(2), 259–270.

Hillis, J. M., Watt, S. J., Landy, M. S. & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of vision*, *4*(12), 1–1.

Hoad, P. (2014). Tetris: How we made the addictive computer game. Retrieved May 28, 2020, from https://www.theguardian.com/culture/2014/jun/02/how-we-made-tetris

Hoffman, D. M., Girshick, A. R., Akeley, K. & Banks, M. S. (2008). Vergence–accommodation conflicts hinder visual performance and cause visual fatigue. *Journal of vision*, *8*(3), 33–33.

Holmboe, K., Larkman, C., de Klerk, C., Simpson, A., Christodoulou, C. & Dvergsdal, H. (2019). The early childhood inhibitory touchscreen task.

Hornsey, R. L. & Hibbard, P. B. (2015). Evaluation of the accuracy of the leap motion controller for measurements of grip aperture, In *Proceedings of the 12th european conference on visual media production*.

Hornsey, R. L., Hibbard, P. B. & Scarfe, P. (2015). Ordinal judgments of depth in monocularly-and stereoscopically-viewed photographs of complex natural scenes, In *2015 international conference on 3d imaging (ic3d)*. IEEE.

Hornsey, R. & Hibbard, P. (2018). Shape and size constancy in consumer virtual reality. *Journal of Vision*, *18*(10), 515–515.

Hornsey, R., Hibbard, P. & Hunter, D. (2017). Distance perception in consumer virtual reality. *Journal of Vision*, *17*(10), 1047–1047.

Hornsey, R., Hibbard, P. & Scarfe, P. (2016). The influence of viewing distance, depth range and inter-camera distance on depth perception and preference judgement-sjudgements in complex natural scenes. *Journal of Vision*, *16*(12), 284–284.

Howard, I. P. & Rogers, B. J. (2012). Perceiving in depth, vol. 2: Stereoscopic vision.

Howard, I. P., Rogers, B. J. Et al. (1995). *Binocular vision and stereopsis*. Oxford University Press, USA.

Howlader, N., Noone, A. M., Krapcho, M., Miller, D., Brest, A., Yu, M., Ruhl, J., Tatalovich, Z., Mariotto, A., Lewis, D. R., Chen, H. S., Feuer, E. J. & Cronin, K. A. (2019). SEER cancer statistics review 1975-2017. *National Cancer Institute*.

Howlett, E. M. (1990). Wide-angle orthostereo, In *Stereoscopic displays and applications*. International Society for Optics and Photonics.

Hu, X. & Hua, H. (2014). Design and assessment of a depth-fused multi-focal-plane display prototype. *Journal of Display Technology*, *10*(4), 308–316.

Hunter, J., Sackier, J. & Berci, G. (1994). Training in laparoscopic cholecystectomy. *Surgical endoscopy*, *8*(1), 28–31.

Hussein, D. M. E.-D. M. (2018). A survey on sentiment analysis challenges. *Journal of King Saud University-Engineering Sciences*, *30*(4), 330–338.

Hyun, M. Y. & O'Keefe, R. M. (2012). Virtual destination image: Testing a telepresence model. *Journal of Business Research*, *65*(1), 29–35.

Iriguchi, M., Koda, H. & Masataka, N. (2018). Colour perception characteristics of women in menopause, In *Proceedings of the 2018 international joint workshop on multimedia artworks analysis and attractiveness computing in multimedia*.

Ivey, C. (2015). Showdown cinematic vr experience released for free! Retrieved May 28, 2020, from https://www.unrealengine.com/en-US/blog/showdown-cinematic-vr-experience-released-for-free

Jack, D., Boian, R., Merians, A. S., Tremaine, M., Burdea, G. C., Adamovich, S. V., Recce, M. & Poizner, H. (2001). Virtual reality-enhanced stroke rehabilitation.

IEEE transactions on neural systems and rehabilitation engineering, *9*(3), 308–318.

Jang, J., Ko, Y., Shin, W. S. & Han, I. (2021). Augmented reality and virtual reality for learning: An examination using an extended technology acceptance model. *IEEE Access*, *9*, 6798–6809.

Jarvis, L. (2019). *Immersive embodiment: Theatres of mislocalized sensation*. Springer Nature.

Jenkin, N. (1957). Effects of varied distance on short-range size judgments. *Journal of experimental psychology*, *54*(5), 327.

Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T. & Walton, A. (2008). Measuring and defining the experience of immersion in games. *International journal of human-computer studies*, *66*(9), 641–661.

Jerome, C. & Witmer, B. (2005). The perception and estimation of egocentric distance in real and augmented reality environments, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision research*, *31*(7-8), 1351–1360.

Jonas, C., Spiller, M. J. & Hibbard, P. (2017). Summation of visual attributes in auditory–visual crossmodal correspondences. *Psychonomic bulletin & review*, *24*(4), 1104–1112.

Jurgens, V., Cockburn, A. & Billinghurst, M. (2006). Depth cues for augmented reality stakeout, In *Proceedings of the 7th acm sigchi new zealand chapter's international conference on computer-human interaction: Design centered hci*.

Kandalaft, M. R., Didehbani, N., Krawczyk, D. C., Allen, T. T. & Chapman, S. B. (2013). Virtual reality social cognition training for young adults with high-functioning autism. *Journal of autism and developmental disorders*, *43*(1), 34–44.

Karhunen, J. (2020). *Effects of camera height on perception of eye height and sense of immersion in omnidirectional images and videos* (Master's thesis). Itä-Suomen yliopisto.

Keefe, B. D., Hibbard, P. B. & Watt, S. J. (2011). Depth-cue integration in grasp programming: No evidence for a binocular specialism. *Neuropsychologia*, *49*(5), 1246–1257.

Keefe, B. D. & Watt, S. J. (2009). The role of binocular vision in grasping: A small stimulus-set distorts results. *Experimental Brain Research*, *194*(3), 435.

Kellogg, S. (2014). *Transformational chairwork: Using psychotherapeutic dialogues in clinical practice*. Rowman & Littlefield.

Kerrigan, I. S. & Adams, W. J. (2013). Learning different light prior distributions for different contexts. *Cognition*, *127*(1), 99–104.

Kersten, D., Mamassian, P. & Knill, D. C. (1997). Moving cast shadows induce apparent motion in depth. *Perception*, *26*(2), 171–192.

Keskinen, T., Mäkelä, V., Kallioniemi, P., Hakulinen, J., Karhu, J., Ronkainen, K., Mäkelä, J. & Turunen, M. (2019). The effect of camera height, actor behavior, and viewer position on the user experience of 360 videos, In *2019 ieee conference on virtual reality and 3d user interfaces (vr)*. IEEE.

Kim, W. S., Ellis, S. R., Tyler, M. E., Hannaford, B. & Stark, L. W. (1987). Quantitative evaluation of perspective and stereoscopic displays in three-axis manual tracking tasks. *IEEE Transactions on Systems, Man, and Cybernetics*, *17*(1), 61–72.

King, W. R. & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & management*, *43*(6), 740–755.

Kirkley, S. (2004). Augmented reality performance assessment battery (arpab): Object recognition, distance estimation and size estimation using optical see-through head-worn displays.

Klein, E., Swan, J. E., Schmidt, G. S., Livingston, M. A. & Staadt, O. G. (2009). Measurement protocols for medium-field distance perception in large-screen immersive displays, In *Virtual reality conference, 2009. vr 2009. ieee*. IEEE.

Klein, L. R. (2003). Creating virtual product experiences: The role of telepresence. *Journal of interactive Marketing*, *17*(1), 41–55.

Kline, P. B. & Witmer, B. G. (1996). Distance perception in virtual environments: Effects of field of view and surface texture at near distances, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Knapp, J. L. J. (2003). Visual perception of egocentric distance in real and virtual environments, In *Virtual and adaptive environments*. CRC Press.

Knapp, J. M. & Loomis, J. M. (2004). Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoperators & Virtual Environments*, *13*(5), 572–577.

Knight, S. (2019). Oculus founder palmer luckey says rift s is only suitable for about 70 percent of the population. Retrieved December 16, 2021, from https://www.techspot.com/news/79345-oculus-founder-palmer-luckey-rift-s-only-suitable.html

Koenderink, J. J. & Van Doorn, A. J. (1995). Relief: Pictorial and otherwise. *Image and Vision Computing*, *13*(5), 321–334.

Koenderink, J. J., Van Doorn, A. J., Kappers, A. M. & Todd, J. T. (2001). Ambiguity and the 'mental eye'in pictorial relief. *Perception*, *30*(4), 431–448.

Koenderink, J. J. & van Doorn, A. J. (1987). Facts on optic flow. *Biological cybernetics*, *56*(4), 247–254.

Koenderink, J. J., van Doorn, A. J., Kappers, A. M., Te Pas, S. F. & Pont, S. C. (2003). Illumination direction from texture shading. *JOSA A*, *20*(6), 987–995.

Koenderink, J. J., van Doorn, A. J. & Pont, S. C. (2004). Light direction from shad (ow) ed random gaussian surfaces. *Perception*, *33*(12), 1405–1420.

Koenderink, J. (1999). Virtual psychophysics. *Perception*, *28*(6), 669–674.

Kopiske, K. K., Bozzacchi, C., Volcic, R. & Domini, F. (2019). Multiple distance cues do not prevent systematic biases in reach to grasp movements. *Psychological research*, *83*(1), 147–158.

Körding, K. P., Beierholm, U., Ma, W. J., Quartz, S., Tenenbaum, J. B. & Shams, L. (2007). Causal inference in multisensory perception. *PLoS one*, *2*(9), e943.

Kramida, G. (2015). Resolving the vergence-accommodation conflict in head-mounted displays. *IEEE transactions on visualization and computer graphics*, *22*(7), 1912–1931.

Krijn, M., Emmelkamp, P. M., Biemond, R., de Ligny, C. d. W., Schuemie, M. J. & van der Mast, C. A. (2004). Treatment of acrophobia in virtual reality: The role of immersion and presence. *Behaviour research and therapy*, *42*(2), 229–239.

Kuhl, S. A., Creem-Regehr, S. H. & Thompson, W. B. (2006). Individual differences in accuracy of blind walking to targets on the floor. *Journal of Vision*, *6*(6), 726–726.

Kumar, G. & Sharma, P. (2014). Google glasses impediments. *medicine*, *1*(2).

Kunnapas, T. (1968). Distance perception as a function of available visual cues. *Journal of experimental psychology*, *77*(4), 523.

Kušen, E. & Strembeck, M. (2018). Politics, sentiments, and misinformation: An analysis of the twitter discussion on the 2016 austrian presidential elections. *Online Social Networks and Media*, *5*, 37–50.

Kytö, M., Mäkinen, A., Häkkinen, J. & Oittinen, P. (2013). Improving relative depth judgments in augmented reality with auxiliary augmentations. *ACM Transactions on Applied Perception (TAP)*, *10*(1), 1–21.

Lampton, D. R., McDonald, D. P., Singer, M. & Bliss, J. P. (1995). Distance estimation in virtual environments, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Landy, M. S., Kording, K. & Trommershauser, J. (2011). *Sensory cue integration*. Oxford University Press.

Lappin, J. S., Shelton, A. L. & Rieser, J. J. (2006). Environmental context influences visually perceived distance. *Perception & psychophysics*, *68*(4), 571–581.

Latané, B. & Rodin, J. (1969). A lady in distress: Inhibiting effects of friends and strangers on bystander intervention. *Journal of Experimental Social Psychology*, *5*(2), 189–202.

Laver, K., George, S., Thomas, S., Deutsch, J. E. & Crotty, M. (2012). Virtual reality for stroke rehabilitation. *Stroke*, *43*(2), e20–e21.

Lee, J., Kim, J. & Choi, J. Y. (2019). The adoption of virtual reality devices: The technology acceptance model integrating enjoyment, social interaction, and strength of the social ties. *Telematics and Informatics*, *39*, 37–48.

Lee, K. M. (2004). Presence, explicated. *Communication theory*, *14*(1), 27–50.

Legris, P., Ingham, J. & Collerette, P. (2003). Why do people use information technology? a critical review of the technology acceptance model. *Information & management*, *40*(3), 191–204.

Leonard, B. (1992). The lawnmower man.

Leue, M., Jung, T. & tom Dieck, D. (2014). A theoretical model of augmented reality acceptance. *E-review of Tourism Research*, *5*.

Leyrer, M., Linkenauger, S. A., Bülthoff, H. H., Kloos, U. & Mohler, B. (2011). The influence of eye height and avatars on egocentric distance estimates in immersive virtual environments, In *Proceedings of the acm siggraph symposium on applied perception in graphics and visualization*.

Leyrer, M., Linkenauger, S. A., Bülthoff, H. H. & Mohler, B. J. (2015). The importance of postural cues for determining eye height in immersive virtual reality. *PloS one*, *10*(5), e0127000.

Li, Z., Phillips, J. & Durgin, F. H. (2011). The underestimation of egocentric distance: Evidence from frontal matching tasks. *Attention, Perception, & Psychophysics*, *73*(7), 2205.

Linev Adani. (2021). *Conpass smart dv*. Retrieved October 27, 2021, from https://security.adanisystems.com/products/adani-people-screening/conpass-smart-dv/

Linton, P. (2020). Does vision extract absolute distance from vergence? *Attention, Perception, & Psychophysics*.

L'Italien, R. (2021). Unreal engine 5 (ue5) has changed the game. Retrieved November 3, 2021, from https://www.perforce.com/blog/vcs/unreal-engine-5

Liu, S. & Hua, H. (2010). A systematic method for designing depth-fused multi-focal plane three-dimensional displays. *Optics express*, *18*(11), 11562–11573.

Livingston, M. A., Ai, Z., Swan, J. E. & Smallman, H. S. (2009). Indoor vs. outdoor depth perception for mobile augmented reality, In *2009 ieee virtual reality conference*. IEEE.

Lombard, M. & Jones, M. T. (2015). Defining presence, In *Immersed in media*. Springer.

Loomis, J. M., Da Silva, J. A., Fujita, N. & Fukusima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 906.

Loomis, J. M., Knapp, J. M. Et al. (2003). Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments*, *11*, 21–46.

Loomis, J. M. & Philbeck, J. W. (1999). Is the anisotropy of perceived 3-d shape invariant across scale? *Perception & psychophysics*, *61*(3), 397–402.

Loomis, J. M. & Philbeck, J. W. (2008). Measuring spatial perception with spatial updating and action., In *Carnegie symposium on cognition, 2006, pittsburgh, pa, us*. Psychology Press.

Loomis, J. M., Philbeck, J. W. & Zahorik, P. (2002). Dissociation between location and shape in visual space. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(5), 1202.

Maani, C. V., Hoffman, H. G., Morrow, M., Maiers, A., Gaylord, K., McGhee, L. L. & DeSocio, P. A. (2011). Virtual reality pain control during burn wound debridement of combat-related burn injuries using robot-like arm mounted vr goggles. *The Journal of trauma*, *71*(1 0), S125.

MacLeod, C. (1999). Anxiety and anxiety disorders. *Handbook of cognition and emotion*, 447–477.

Mamassian, P. (2004). Impossible shadows and the shadow correspondence problem. *Perception*, *33*(11), 1279–1290.

Mamassian, P., Knill, D. C. & Kersten, D. (1998). The perception of cast shadows. *Trends in cognitive sciences*, *2*(8), 288–295.

Manguri, K. H., Ramadhan, R. N. & Amin, P. R. M. (2020). Twitter sentiment analysis on worldwide covid-19 outbreaks. *Kurdistan Journal of Applied Research*, 54–65.

Manis, K. T. & Choi, D. (2019). The virtual reality hardware acceptance model (vr-ham): Extending and individuating the technology acceptance model (tam) for virtual reality hardware. *Journal of Business Research*, *100*, 503–513.

Masaoka, K., Hanazato, A., Emoto, M., Yamanoue, H., Nojiri, Y. & Okano, F. (2006). Spatial distortion prediction system for stereoscopic images. *Journal of Electronic Imaging*, *15*(1), 013002.

MasterClass. (2021). Film 101: Understanding film lighting. Retrieved May 28, 2021, from https://www.masterclass.com/articles/film-101-understanding-film-lighting#what-is-cinematic-lighting

McKeand, K. (2019). One of the reasons borderlands 3 uses unreal engine is because it's compatible with next-gen. Retrieved May 28, 2020, from https://www.vg247.com/2019/08/14/borderlands-3-next-gen/

McKenna, K., McMenemy, K., Ferguson, R. S., Dick, A. & Potts, S. (2008). Enhancing the immersive reality of virtual simulators for easily accessible laparoscopic surgical training, In *The engineering reality of virtual reality 2008*. International Society for Optics and Photonics.

Meehan, M., Razzaque, S., Whitton, M. C. & Brooks, F. P. (2003). Effect of latency on presence in stressful virtual environments, In *Ieee virtual reality, 2003. proceedings.* IEEE.

Mercier, O., Sulai, Y., Mackenzie, K., Zannoli, M., Hillis, J., Nowrouzezahrai, D. & Lanman, D. (2017). Fast gaze-contingent optimal decompositions for multifocal displays. *ACM Transactions on Graphics (TOG)*, *36*(6), 237.

Messing, R. & Durgin, F. H. (2005). Distance perception and the visual horizon in head-mounted displays. *ACM Transactions on Applied Perception (TAP)*, *2*(3), 234–250.

Meyerhoff, H. S. & Papenmeier, F. (2020). Individual differences in visual attention: A short, reliable, open source, and multilingual test of multiple object tracking in psychopy. *PsyArXiv. June*, *10*.

Miglicco, G. (2018). Gdpr is here and it is time to get serious. *Computer Fraud & Security*, *2018*(9), 9–12.

MINDFLUX. (2006). Mindflux about. Retrieved July 19, 2021, from http://www.mindflux.com.au/about.html

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Group, P. Et al. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS medicine*, *6*(7), e1000097.

Mohler, B. J., Bülthoff, H. H., Thompson, W. B. & Creem-Regehr, S. H. (2008). A full-body avatar improves egocentric distance judgments in an immersive virtual environment, In *Proceedings of the 5th symposium on applied perception in graphics and visualization*.

Mohler, B. J., Creem-Regehr, S. H. & Thompson, W. B. (2006). The influence of feedback on egocentric distance judgments in real and virtual environments, In *Proceedings of the 3rd symposium on applied perception in graphics and visualization*. ACM.

Monstars Inc. and Resonair. (2019). *Tetris effect* [Intrance, Inc]. Retrieved May 28, 2020, from https://www.epicgames.com/store/en-US/product/tetris-effect/home

Mon-Williams, M., Wann, J. & Rushton, S. (1995). Design factors in stereoscopic virtual-reality displays. *Journal of the Society for information Display*, *3*(4), 207–210.

Mon-Williams, M. & Tresilian, J. R. (1999). Some recent studies on the extraretinal contribution to distance perception. *Perception*, *28*(2), 167–181.

Mon-Williams, M. & Tresilian, J. R. (2000). Ordinal depth information from accommodation? *Ergonomics*, *43*(3), 391–404.

Mon-Williams, M., Tresilian, J. R. & Roberts, A. (2000). Vergence provides veridical depth perception from horizontal retinal image disparities. *Experimental brain research*, *133*(3), 407–413.

Morrison, J. D. & Whiteside, T. C. (1984). Binocular cues in the perception of distance of a point source of light. *Perception*, *13*(5), 555–566.

Moyon, A., Poirson, E. & Petiot, J.-F. (2019). Development of an acceptance model for occupational exoskeletons and application for a passive upper limb device. *IISE Transactions on Occupational Ergonomics and Human Factors*, *7*(3-4), 291–301.

Murgia, A., Sharkey, P. M. Et al. (2009). Estimation of distances in virtual environments using size constancy. *The International Journal of Virtual Reality*, *8*(1), 67–74.

Murray, R. F. & Morgenstern, Y. (2010). Cue combination on the circle and the sphere. *Journal of vision*, *10*(11), 15–15.

Nagase, M., Kimoto, Y., Sunami, E. & Matsumura, G. (2020). A new human cadaver model for laparoscopic training using n-vinyl-2-pyrrolidone: A feasibility study. *Anatomical science international*, *95*(1), 156–164.

Nagata, S. (1991). How to reinforce perception of depth in single two-dimensional pictures. *Pictorial communication in virtual and real environments*.

Nakano, G., Kitahara, I. & Ohta, Y. (2008). Generating perceptually-correct shadows for mixed reality, In *2008 7th ieee/acm international symposium on mixed and augmented reality*. IEEE.

Napieralski, P. E., Altenhoff, B. M., Bertrand, J. W., Long, L. O., Babu, S. V., Pagano, C. C., Kern, J. & Davis, T. A. (2011). Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Transactions on Applied Perception (TAP)*, *8*(3), 1–19.

Narciso, D., Melo, M., Raposo, J. V., Cunha, J. & Bessa, M. (2020). Virtual reality in training: An experimental study with firefighters. *Multimedia Tools and Applications*, *79*(9), 6227–6245.

Neale, D. (1998). Head-mounted displays: Product reviews and related design considerations. *Considerations, Hypermedia Technical Report HCIL-98-02, HumanComputer Interaction Laboratory, Department of Industrial and Systems Engineering, Virginia Tech*.

Netflix, Inc. (2018). Netflix vr. Retrieved December 15, 2021, from https://play.google.com/store/apps/details?id=com.netflix.android_vr%5C&hl=en_GB%5C&gl=US

NetherRealm Studios. (2015). *Mortal kombat x* [Warner Bros. Interactive Entertainment].

Nicolaou, M., James, A., Lo, B. P., Darzi, A. & Yang, G.-Z. (2005). Invisible shadow for navigation and planning in minimal invasive surgery, In *International conference on medical image computing and computer-assisted intervention*. Springer.

Niehorster, D. C., Li, L. & Lappe, M. (2017). The accuracy and precision of position and orientation tracking in the htc vive virtual reality system for scientific research. *i-Perception*, *8*(3), 2041669517708205.

Nielsen, R. S. (2003). *Real time rendering of— atmospheric scattering effects for flight simulators* (Master's thesis). Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark.

Oculus. (2012). Oculus device specifications. Retrieved July 20, 2021, from https://developer.oculus.com/resources/oculus-device-specs/

Oculus. (2017). Oculus rift — step into rift – now only $ 399. https://www.youtube.com/watch?v=5q6BcQq%7B_%7Dyhw

Oculus. (2021). Quest 2. Retrieved November 4, 2021, from https://www.oculus.com/quest-2/?locale=en-gb%5C&utm_source=gg%5C&utm_medium=pla%5C&utm_campaign=%2011139634546%5C&utm_term%5C&utm_content=465631893315%5C&utm_ad=108872177213%5C&%20utm_location=1006644%5C&utm_location2%5C&utm_placement=pla-1485608734207%5C&%20utm_device=c%5C&utm_matchtype%5C&utm_feed%5C&utm_adposi%20tion%5C&utm_product=899-%2000187-02%5C&gclsrc=aw.ds%5C&%5C&gclid=CjwKCAjwiY6%20MBhBq%20EiwARFSCPmGVvfF4%20IkOLNCns9bbN5y8ah%20Fqv07YqXneAg%20O%20RJzZDIgNsYanPiBRoCwZEQAvD_BwE

Oh, C. S., Bailenson, J. N. & Welch, G. F. (2018). A systematic review of social presence: Definition, antecedents, and implications. *Frontiers in Robotics and AI*, *5*, 114.

O'Hare, L., Zhang, T., Nefs, H. T. & Hibbard, P. B. (2013). Visual discomfort and depth-of-field. *i-Perception*, *4*(3), 156–169.

Olivi, G. (2019). Virtual reality: Top data protection issues to consider. Retrieved November 12, 2021, from https://www.dentons.com/en/insights/articles/2019/october/18/virtual-reality-top-data-protection-issues-to-consider

Ooi, T. L., Wu, B. & He, Z. J. (2001). Distance determined by the angular declination below the horizon. *Nature*, *414*(6860), 197–200.

O'regan, J. K. & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and brain sciences*, *24*(5), 939–973.

Oruç, I., Maloney, L. T. & Landy, M. S. (2003). Weighted linear cue combination with possibly correlated error. *Vision research*, *43*(23), 2451–2468.

Owsley, C. & Sloane, M. E. (1987). Contrast sensitivity, acuity, and the perception of 'real-world' targets. *British Journal of Ophthalmology*, *71*(10), 791–796.

Paige, G. D., Telford, L., Seidman, S. H. & Barnes, G. R. (1998). Human vestibuloocular reflex and its interactions with vision and fixation distance during linear and angular head movement. *Journal of Neurophysiology*, *80*(5), 2391–2404.

Paillard, J. (1991). Motor and representational framing of space. *Brain and space*, 163–182.

Palmisano, S., Gillam, B., Govan, D. G., Allison, R. S. & Harris, J. M. (2010). Stereoscopic perception of real depths at large distances. *Journal of vision*, *10*(6), 19–19.

Pan, X. & Hamilton, A. F. D. C. (2018). Why and how to use virtual reality to study human social interaction: The challenges of exploring a new research landscape. *British Journal of Psychology*, *109*(3), 395–417.

Parsons, S. & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of intellectual disability research*, *46*(5), 430–443.

Parsons, T. D. & Rizzo, A. A. (2008). Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *Journal of behavior therapy and experimental psychiatry*, *39*(3), 250–261.

Patchs. (2016). *Food pack 01* [Unreal Engine Marketplace]. Retrieved September 14, 2020, from https://www.unrealengine.com/marketplace/en-US/product/food-pack

Patterson, R., Moe, L. & Hewitt, T. (1992). Factors that affect depth perception in stereoscopic displays. *Human factors*, *34*(6), 655–667.

Peirce, J. W. (2007). Psychopy—psychophysics software in python. *Journal of neuroscience methods*, *162*(1-2), 8–13.

Pelli, D. G. & Vision, S. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial vision*, *10*, 437–442.

Perkins Coie. (2020). Industry insights into the future of immersive technology. *2020 AUGMENTED AND VIRTUAL REALITY SURVEY REPORT*, *4*. https://www.perkinscoie.com/images/content/2/3/v4/231654/2020-AR-VR-Survey-v3.pdf#page=23%5C&zoom=100,0,0

Petty, R. E., Wegener, D. T. & Fabrigar, L. R. (1997). Attitudes and attitude change. *Annual review of psychology*, *48*(1), 609–647.

Plumert, J. M., Kearney, J. K., Cremer, J. F. & Recker, K. (2005). Distance perception in real and virtual environments. *ACM Transactions on Applied Perception (TAP)*, *2*(3), 216–233.

Ponto, K., Gleicher, M., Radwin, R. G. & Shin, H. J. (2013). Perceptual calibration for immersive display environments. *IEEE transactions on visualization and computer graphics*, *19*(4), 691–700.

Proffitt, D. R. (2006). Distance perception. *Current Directions in psychological science*, *15*(3), 131–135.

Proffitt, D. R. & Caudek, C. (2003). Depth perception and the perception of events. *Handbook of psychology*, 213–236.

Prothero, J. D., Hoffman, H. G., Parker, D. E., Furness III, T. A. & Wells, M. J. (1995). Foreground/background manipulations affect presence, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Purdy, J. & Gibson, E. J. (1955). Distance judgment by the method of fractionation. *Journal of Experimental Psychology*, *50*(6), 374.

Ratzel, M. (2020). *Concept study for the use of a render engine for computer games in a flight simulation* (Doctoral dissertation). DLR-Institut für Flugsystemtechnik.

RDocumentation. (2021a). Mcnemar.test: Mcnemar's chi-squared test for count data [CRAN]. Retrieved October 21, 2021, from https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/mcnemar.test

RDocumentation. (2021b). Wilcox.test: Wilcoxon rank sum and signed rank tests [CRAN]. Retrieved October 14, 2021, from https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/wilcox.test

Renner, R. S., Velichkovsky, B. M. & Helmert, J. R. (2013). The perception of egocentric distances in virtual environments-a review. *ACM Computing Surveys (CSUR)*, *46*(2), 23.

Richards, W. & Miller, J. F. (1969). Convergence as a cue to depth. *Perception & Psychophysics*, *5*(5), 317–320.

Rieser, J. J., Ashmead, D. H., Talor, C. R. & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, *19*(5), 675–689.

Robinson, M. (2018). Unreal engine fuels "destination imagination" virtual reality immersive experience at SCAD. Retrieved May 28, 2020, from https://www.unrealengine.com/en-US/developer-interviews/unreal-engine-fuels-destination-imagination-virtual-reality-immersive-experience-at-scad

Roddenberry, G. (1987). Star trek: The next generation.

Rogers, B. (2011). Information, illusion and constancy in telestereoscopic viewing, In *Harris, l and jenkin m (eds)"perception of 3-d environments"*. Cambridge University Press.

Rohde, M., Van Dam, L. C. & Ernst, M. O. (2014). Predictability is necessary for closed-loop visual feedback delay adaptation. *Journal of vision*, *14*(3), 4–4.

Rohde, M., van Dam, L. C. & Ernst, M. O. (2016). Statistically optimal multisensory cue integration: A practical tutorial. *Multisensory research*, *29*(4-5), 279–317.

Roitman, J. D. & Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *Journal of neuroscience*, *22*(21), 9475–9489.

Rolland, J. P., Meyer, C., Arthur, K. & Rinalducci, E. (2002). Method of adjustments versus method of constant stimuli in the quantification of accuracy and precision of rendered depth in head-mounted displays. *Presence*, *11*(6), 610–625.

Roumes, C., Meehan, J. W., Plantier, J. & Menu, J.-P. (2001). Distance estimation in a 3-d imaging display. *The International Journal of Aviation Psychology*, *11*(4), 381–396.

Safran, J. D., Muran, J. C., Stevens, C. & Rothman, M. (2008). A relational approach to supervision: Addressing ruptures in the alliance.

Sagnier, C., Loup-Escande, E., Lourdeaux, D., Thouvenin, I. & Valléry, G. (2020). User acceptance of virtual reality: An extended technology acceptance model. *International Journal of Human–Computer Interaction*, *36*(11), 993–1007.

Sakurai, S., Narumi, T., Ban, Y., Tanikawa, T. & Hirose, M. (2013). Affecting our perception of satiety by changing the size of virtual dishes displayed with a tabletop display, In *International conference on virtual, augmented and mixed reality*. Springer.

Samsung. (2015). Gear vr specifications. Retrieved July 20, 2021, from https://www.samsung.com/global/galaxy/gear-vr/specs/

Samsung. (2021a). Galaxy s21 ultra 5g specs. Retrieved July 20, 2021, from https://www.samsung.com/uk/smartphones/galaxy-s21-ultra-5g/specs/

Samsung. (2021b). Galaxya52 5g specs. Retrieved July 20, 2021, from https://www.samsung.com/uk/smartphones/galaxy-a/galaxy-a52-5g-awesome-black-128gb-sm-a526bzkdeua/

Sanchez-Vives, M. V. & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews Neuroscience*, *6*(4), 332–339.

Sarupuri, B., Chipana, M. L. & Lindeman, R. W. (2017). Trigger walking: A low-fatigue travel technique for immersive virtual reality, In *2017 ieee symposium on 3d user interfaces (3dui)*. IEEE.

Scarfe, P. & Glennerster, A. (2015). Using high-fidelity virtual reality to study perception in freely moving observers. *Journal of vision*, *15*(9), 3–3.

Scarfe, P. & Glennerster, A. (2019). The science behind virtual reality displays. *Annual review of vision science*, *5*, 529–547.

Scarfe, P. & Hibbard, P. B. (2011). Statistically optimal integration of biased sensory estimates. *Journal of vision*, *11*(7), 12–12.

Scarfe, P. & Hibbard, P. B. (2013). Reverse correlation reveals how observers sample visual information when estimating three-dimensional shape. *Vision Research*, *86*, 115–127.

Scarfe, P. & Hibbard, P. (2006). Disparity-defined objects moving in depth do not elicit three-dimensional shape constancy. *Vision research*, *46*(10), 1599–1610.

Schwind, V., Knierim, P., Haas, N. & Henze, N. (2019). Using presence questionnaires in virtual reality, In *Proceedings of the 2019 chi conference on human factors in computing systems*.

Sciuto, L. A. L. & Hartley, E. L. (1963). Religious affiliation and open-mindedness in binocular resolution. *Perceptual and motor skills*, *17*(2), 427–430.

Sedgwick, H. A. (1986). Space perception. *Sensory proces and perception*.

Segment. (2017). The 2017 state of personalization report. Retrieved October 21, 2021, from http://grow.segment.com/Segment-2017-Personalization-Report.pdf

Seymour, N. E., Gallagher, A. G., Roman, S. A., O'Brien, M. K., Bansal, V. K., Andersen, D. K. & Satava, R. M. (2002). Virtual reality training improves operating room performance: Results of a randomized, double-blinded study. *Annals of surgery*, *236*(4), 458.

Sharman, R. J., McGraw, P. V. & Peirce, J. W. (2013). Luminance cues constrain chromatic blur discrimination in natural scene stimuli. *Journal of Vision*, *13*(4), 14–14.

Sharman, R. J., McGraw, P. V. & Peirce, J. W. (2015). Cue combination of conflicting color and luminance edges. *i-Perception*, *6*(6), 2041669515621215.

Shibata, T., Kim, J., Hoffman, D. M. & Banks, M. S. (2011a). Visual discomfort with stereo displays: Effects of viewing distance and direction of vergence-accommodation conflict, In *Stereoscopic displays and applications xxii*. International Society for Optics and Photonics.

Shibata, T., Kim, J., Hoffman, D. M. & Banks, M. S. (2011b). The zone of comfort: Predicting visual discomfort with stereo displays. *Journal of vision*, *11*(8), 11–11.

Shimi, A., Tsestou, V., Hadjiaros, M., Neokleous, K. & Avraamides, M. (2021). Attentional skills in soccer: Evaluating the involvement of attention in executing a goalkeeping task in virtual reality. *Applied Sciences*, *11*(19), 9341.

Shimotsu, R. T. & Cao, C. G. (2007). The effect of color-contrasting shadows on a dynamic 3-d laparoscopic surgical task. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *37*(6), 1047–1053.

Siegel, M. & Nagata, S. (2000). Just enough reality: Comfortable 3-d viewing via microstereopsis. *IEEE Transactions on Circuits and Systems for Video Technology*, *10*(3), 387–396.

Simon, T., Pagel, S. & von Korflesch, H. F. O. (2020). Influencing factors for acceptance of digital tools in the humanities, In *Proceedings of the conference on mensch und computer*, Magdeburg, Germany, Association for Computing Machinery.

Simons, D. J., Franconeri, S. L. & Reimer, R. L. (2000). Change blindness in the absence of a visual disruption. *Perception*, *29*(10), 1143–1154.

Sinai, M. J., Krebs, W. K., Darken, R. P., Rowland, J. & McCarley, J. (1999). Egocentric distance perception in a virutal environment using a perceptual matching task, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Sinai, M. J., Ooi, T. L. & He, Z. J. (1998). Terrain influences the accurate judgement of distance. *Nature*, *395*(6701), 497.

Singh, G., Swan II, J. E., Jones, J. A. & Ellis, S. R. (2010). Depth judgement measures and occluding surfaces in near-field augmented reality, In *Proceedings of the 7th symposium on applied perception in graphics and visualization*. ACM.

Singh, R. P., Javaid, M., Kataria, R., Tyagi, M., Haleem, A. & Suman, R. (2020). Significant applications of virtual reality for covid-19 pandemic. *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, *14*(4), 661–664.

Slater, M. (2003). A note on presence terminology. *Presence connect*, *3*(3), 1–5.

Slater, M. (2004). How colorful was your day? why questionnaires cannot assess presence in virtual environments. *Presence*, *13*(4), 484–493.

Slater, M., Linakis, V., Usoh, M. & Kooper, R. (1996). Immersion, presence and performance in virtual environments: An experiment with tri-dimensional chess, In *Proceedings of the acm symposium on virtual reality software and technology*. ACM.

Slater, M. & Steed, A. (2000). A virtual presence counter. *Presence: Teleoperators & Virtual Environments*, *9*(5), 413–434.

Slater, M., Usoh, M. & Steed, A. (1995). Taking steps: The influence of a walking technique on presence in virtual reality. *ACM Transactions on Computer-Human Interaction (TOCHI)*, *2*(3), 201–219.

Smallman, H. S., St. John, M. & Cowen, M. B. (2002). Use and misuse of linear perspective in the perceptual reconstruction of 3-d perspective view displays, In *Proceedings of the human factors and ergonomics society annual meeting*. SAGE Publications Sage CA: Los Angeles, CA.

Smith, M., Doutcheva, N., Gabbard, J. L. & Burnett, G. (2015). Optical see-through head up displays' effect on depth judgments of real world objects, In *2015 ieee virtual reality (vr)*. IEEE.

Sogo, H. (2017). Sgttoolbox: Utility for controlling simplegazetracker from psychtoolbox. *Behavior research methods*, *49*(4), 1323–1332.

Sport England. (2021). Understanding the impact of covid-19. Retrieved October 19, 2021, from https://sportengland-production-files.s3.eu-west-2.amazonaws.com/s3fs-public/2021-04/Understanding%5C%20the%5C%20Impact%5C%20of%5C%20Covid%5C%20April%5C%202021_0.pdf?VersionId=rDJkuKjVEnrsQYsD%20n9nSYezUmXlu6ZK9%5C#page=21%5C&zoom=100,0,0

Sprague, W. W., Cooper, E. A., Reissier, S., Yellapragada, B. & Banks, M. S. (2016). The natural statistics of blur. *Journal of vision*, *16*(10), 23–23.

Star Strider. (2015). How to generate noise with particular frequency? [blog post]. Retrieved June 2, 2020, from https://uk.mathworks.com/matlabcentral/answers/244276-how-to-generate-noise-with-particular-frequency

Steam. (2021). Valve index. Retrieved October 20, 2021, from https://store.steampowered.com/valveindex

Steinfeld, N. (2016). "i agree to the terms and conditions": (how) do users read privacy policies online? an eye-tracking experiment. *Computers in human behavior*, *55*, 992–1000.

Stelter, A., Kaping, C., Oschinsky, F. M. & Niehaves, B. (2020). Theoretical foundations on technology acceptance and usage in public administrations: Investigating bounded acceptance and usage of new technology by employees, In *The 21st annual international conference on digital government research*, Seoul, Republic of Korea, Association for Computing Machinery.

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

Strickland, D. (2017). Borderlands 3 will be powered by unreal engine 4. Retrieved May 28, 2020, from https://www.tweaktown.com/news/56509/borderlands-3-powered-unreal-engine-4/index.html

Stroustrup, B. (2007). Evolving a language in and for the real world: C++ 1991-2006, In *Proceedings of the third acm sigplan conference on history of programming languages*.

Suchotzki, K., Verschuere, B., Van Bockstaele, B., Ben-Shakhar, G. & Crombez, G. (2017). Lying takes time: A meta-analysis on reaction time measures of deception. *Psychological Bulletin*, *143*(4), 428.

Sugano, N., Kato, H. & Tachibana, K. (2003). The effects of shadow representation of virtual objects in augmented reality, In *The second ieee and acm international symposium on mixed and augmented reality, 2003. proceedings.* IEEE.

Surdick, R. T., Davis, E. T., King, R. A. & Hodges, L. F. (1997). The perception of distance in simulated visual displays: A comparison of the effectiveness and accuracy of multiple depth cues across viewing distances. *Presence: Teleoperators & Virtual Environments*, *6*(5), 513–531.

Sutherland, I. E. (1965). The ultimate display. *Multimedia: From Wagner to virtual reality*, 506–508.

Sutherland, I. E. (1968). A head-mounted three dimensional display, In *Proceedings of the december 9-11, 1968, fall joint computer conference, part i*. ACM.

Swan, J. E., Livingston, M. A., Smallman, H. S., Brown, D., Baillot, Y., Gabbard, J. L. & Hix, D. (2006). A perceptual matching technique for depth judgments in optical, see-through augmented reality, In *Ieee virtual reality conference (vr 2006)*. IEEE.

te Pas, S. F., Pont, S. C., Dalmaijer, E. S. & Hooge, I. T. (2017). Perception of object illumination depends on highlights and shadows, not shading. *Journal of vision*, *17*(8), 2–2.

Tharp, G. K., Liu, A. M., French, L., Lai, S. & Stark, L. W. (1992). Timing considerations of helmet-mounted display performance, In *Human vision, visual processing, and digital display iii*. International Society for Optics and Photonics.

Thomas, G., Goldberg, J. H., Cannon, D. J. & Hillis, S. L. (2002). Surface textures improve the robustness of stereoscopic depth cues. *Human factors*, *44*(1), 157–170.

Thomas, N. J. (1999). Are theories of imagery theories of imagination? an active perception approach to conscious mental content. *Cognitive science*, *23*(2), 207–245.

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M. & Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments? *Presence: Teleoperators & Virtual Environments*, *13*(5), 560–571.

Thomsen, M. (2012). History of the unreal engine. Retrieved May 28, 2020, from https://www.ign.com/articles/2010/02/23/history-of-the-unreal-engine

Todd, J. T. (2020). On the ambient optic array: James gibson's insights about the phenomenon of chiaroscuro. *i-Perception*, *11*(5), 2041669520952097.

Todd, J. T., Christensen, J. C. & Guckes, K. M. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture? *Journal of Vision*, *10*(2), 20–20.

Tresilian, J. R., Mon-Williams, M. & Kelly, B. M. (1999). Increasing confidence in vergence as a cue to distance. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1414), 39–44.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *59*(236), 433–460. http://0-www-jstor-org.serlib0.essex.ac.uk/stable/2251299

UbiSim. (2021). *The world's first immersive vr training platform for nursing*. Retrieved September 3, 2021, from https://www.ubisimvr.com/

Ultraleap. (2021). *Grasp the true potential of xr*. Retrieved November 16, 2021, from https://www.ultraleap.com/

Van Dam, L. C. & Stephens, J. R. (2018). Effects of prolonged exposure to feedback delay on the qualitative subjective experience of virtual reality. *PloS one*, *13*(10).

Van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS quarterly*, 695–704.

Van der Meijden, O. A. & Schijven, M. P. (2009). The value of haptic feedback in conventional and robot-assisted minimal invasive surgery and virtual reality training: A current review. *Surgical endoscopy*, *23*(6), 1180–1190.

Van Ee, R., Van Dam, L. C. & Erkelens, C. J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision*, *2*(9), 2–2.

Van Koningsbruggen, G. M., Stroebe, W. & Aarts, H. (2011). Through the eyes of dieters: Biased size perception of food following tempting food primes. *Journal of Experimental Social Psychology*, *47*(2), 293–299.

van Dam, L. C., Webb, A. L., Jarvis, L. D., Hibbard, P. B. & Linley, M. (2020). ''the mystery of the raddlesham mumps": A case study for combined storytelling in a theatre play and virtual reality, In *2020 international conference on 3d immersion (ic3d)*. IEEE.

Vasquez, J. M., Vaca, V. L., Wiederhold, B. K., Miller, I. & Wiederhold, M. D. (2017). Virtual reality pain distraction during gynecological surgery - a report of 44 cases. *Surgical Research Updates*.

Venkatesh, V., Morris, M. G., Davis, G. B. & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425–478.

Venkatesh, V. & Speier, C. (1999). Computer technology training in the workplace: A longitudinal investigation of the effect of mood. *Organizational behavior and human decision processes*, *79*(1), 1–28.

Verma, I. M. (2014). Editorial expression of concern: Experimental evidence of massivescale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 201412469.

Veron, H., Hezel, P. & Southard, D. A. (1994). Head-mounted displays for virtual reality, In *Helmet-and head-mounted displays and symbology design requirements*. International Society for Optics and Photonics.

Viguier, A., Clement, G. & Trotter, Y. (2001). Distance perception within near visual space. *Perception*, *30*(1), 115–124.

Vishwanath, D. (2014). Toward a new theory of stereopsis. *Psychological review*, *121*(2), 151.

Vishwanath, D. & Blaser, E. (2010). Retinal blur and the perception of egocentric distance. *Journal of Vision*, *10*(10), 26–26.

Vishwanath, D. & Hibbard, P. (2013). Seeing in 3d with just one eye: Stereopsis in the absence of binocular disparities. *Psychological Science*, *24*, 1673–1685.

Vive. (2019). Specs & details. Retrieved July 20, 2021, from https://www.vive.com/uk/product/vive/

Vive. (2021a). Adjusting the ipd on the headset. Retrieved December 16, 2021, from https://www.vive.com/uk/support/vive-pro-hmd/category_howto/adjusting-the-ipd-on-the-headset.html

Vive. (2021b). Products. Retrieved October 20, 2021, from https://www.vive.com/uk/product/

Vive. (2021c). Vive wireless adapter. Retrieved November 4, 2021, from https://www.vive.com/uk/accessory/wireless-adapter/

Von Hildebrand, A. (1905). *Das problem der form in der bildenden kunst*. Heitz.

von Castell, C., Oberfeld, D. & Hecht, H. (2021). Visual and postural eye-height information is flexibly coupled in the perception of virtual environments. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(8), 1132.

Wachowski, L. & Wachowski, L. (1999). The matrix.

Waller, D. (1999). Factors affecting the perception of interobject distances in virtual environments. *Presence*, *8*(6), 657–670.

Wallraven, C., Bülthoff, H. H., Waterkamp, S., van Dam, L. & Gaißert, N. (2014). The eyes grasp, the hands see: Metric category knowledge transfers between vision and touch. *Psychonomic bulletin & review*, *21*(4), 976–985.

Wanger, L. R., Ferwerda, J. A. & Greenberg, D. P. (1992). Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, *12*(3), 44–58.

Wann, J. P., Rushton, S. & Mon-Williams, M. (1995). Natural problems for stereoscopic depth perception in virtual environments. *Vision research*, *35*(19), 2731–2736.

Watt, S. J., Akeley, K., Ernst, M. O. & Banks, M. S. (2005a). Focus cues affect perceived depth. *Journal of vision*, *5*(10), 7–7.

Watt, S. J., Akeley, K., Girshick, A. R. & Banks, M. S. (2005b). Achieving near-correct focus cues in a 3d display using multiple image planes, In *Human vision and electronic imaging x*. International Society for Optics and Photonics.

Watt, S. J., MacKenzie, K. J. & Ryan, L. (2012). Real-world stereoscopic performance in multiple-focal-plane displays: How far apart should the image planes be?, In *Stereoscopic displays and applications xxiii*. SPIE.

Watt, S., Keefe, B. & Hibbard, P. (2007). Do binocular depth cues have a special role in grasping? *Journal of Vision*, *7*(9), 163–163.

Wawro, A. (2016). Arkham vr and rocksteady's path to vr game development. Retrieved May 28, 2020, from https://web.archive.org/web/20160922071131/https://www.

gamasutra.com/view/news/275190/Arkham_VR_and_Rocksteadys_path_to_VR_game_development.php

Webb, A. L., Hibbard, P. B. & O'Gorman, R. (2018). Natural variation in female reproductive hormones does not affect contrast sensitivity. *Royal Society open science*, *5*(2), 171566.

Welchman, A. E., Deubelius, A., Conrad, V., Bulthoff, H. H. & Kourtzi, Z. (2005). 3d shape perception from combined depth cues in human visual cortex. *Nature neuroscience*, *8*(6), 820.

Werkhoven, P. J. & Groen, J. (1998). Manipulation performance in interactive virtual environments. *Human Factors*, *40*(3), 432–442.

Whitehouse, J. & Gruber, H. (1957). The effects of surface-texture and binocular disparity on distance-ratio perception. *Colorado Psychol. Assn, Fort Collins*.

Wickens, C. D. (1990). Three-dimensional stereoscopic display implementation: Guidelines derived from human visual capabilities, In *Stereoscopic displays and applications*. International Society for Optics and Photonics.

Willemsen, P., Colton, M. B., Creem-Regehr, S. H. & Thompson, W. B. (2004). The effects of head-mounted display mechanics on distance judgments in virtual environments, In *Proceedings of the 1st symposium on applied perception in graphics and visualization*.

Willemsen, P., Gooch, A. A., Thompson, W. B. & Creem-Regehr, S. H. (2008). Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoperators and Virtual Environments*, *17*(1), 91–101.

Witmer, B. G. & Kline, P. B. (1998). Judging perceived and traversed distance in virtual environments. *Presence*, *7*(2), 144–167.

Witmer, B. G. & Sadowski Jr, W. J. (1998). Nonvisually guided locomotion to a previously viewed target in real and virtual environments. *Human factors*, *40*(3), 478–488.

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

Wohlwill, J. F. (1963). Overconstancy in distance perception as a function of the texture of the stimulus field and other variables. *Perceptual and Motor Skills*, *17*(3), 831–846.

Woods, A. J., Docherty, T. & Koch, R. (1993). Image distortions in stereoscopic video systems, In *Stereoscopic displays and applications iv*. International Society for Optics and Photonics.

Wu, B., Ooi, T. L. & He, Z. J. (2004). Perceiving distance accurately by a directional process of integrating ground information. *Nature*, *428*(6978), 73–77.

Wunderlich, F. & Memmert, D. (2020). Innovative approaches in sports science—lexicon-based sentiment analysis as a tool to analyze sports-related twitter communication. *Applied Sciences*, *10*(2), 431.

Xia, L., Pont, S. C. & Heynderickx, I. (2014). The visual light field in real scenes. *i-Perception*, *5*(7), 613–629.

Yan, Y., Yu, C., Ma, X., Huang, S., Iqbal, H. & Shi, Y. (2018). Eyes-free target acquisition in interaction space around the body for virtual reality, In *Proceedings of the 2018 chi conference on human factors in computing systems*.

Yang, S. & Sheedy, J. E. (2011). Effects of vergence and accommodative responses on viewer's comfort in viewing 3d stimuli, In *Stereoscopic displays and applications xxii*. International Society for Optics and Photonics.

Yang, Z. & Purves, D. (2003). A statistical explanation of visual space. *Nature neuroscience*, *6*(6), 632.

Yessenalina, A., Yue, Y. & Cardie, C. (2010). Multi-level structured models for document-level sentiment classification, In *Proceedings of the 2010 conference on empirical methods in natural language processing*.

Yonas, A., Goldsmith, L. T. & Hallstrom, J. L. (1978). Development of sensitivity to information provided by cast shadows in pictures. *Perception*, *7*(3), 333–341.

Young, M. J., Landy, M. S. & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision research*, *33*(18), 2685–2696.

Yu, W., Wen, L., Zhao, L.-A., Liu, X., Wang, B. & Yang, H. (2019). The applications of virtual reality technology in medical education: A review and mini-research, In *Journal of physics: Conference series*. IOP Publishing.

Yuan, Y. & Steed, A. (2010). Is the rubber hand illusion induced by immersive virtual reality?, In *2010 ieee virtual reality conference (vr)*. IEEE.

Zahorik, P. (2002). Assessing auditory distance perception using virtual acoustics. *The Journal of the Acoustical Society of America*, *111*(4), 1832–1846.

Zahorik, P., Brungart, D. S. & Bronkhorst, A. W. (2005). Auditory distance perception in humans: A summary of past and present research. *ACTA Acustica united with Acustica*, *91*(3), 409–420.

Zhang, R. & Hua, H. (2010). Effects of a retroreflective screen on depth perception in a head-mounted projection display, In *2010 ieee international symposium on mixed and augmented reality*. IEEE.

Zhong, F., Jindal, A., Yöntem, Ö., Hanji, P., Watt, S. & Mantiuk, R. (2021). Reproducing reality with a high-dynamic-range multi-focal stereo display. *ACM Transactions on Graphics*, *40*(6), 241.