PERCEPTUAL ISSUES OF VISUAL ATTENTION AND DEPTH PERCEPTION IN AUGMENTED REALITY

by

Michael Long, BSc, MSc

Thesis submitted to the University of Nottingham

For the degree of Doctor of Philosophy

January 2018

Thesis Supervisor: Dr Harriet Allen

Table of contents

Chapter 1. Introduction.....1

1.1Research Motivation 1

1.1.1 Jaguar Land Rover.....1

1.1.2 Theoretical motivation 2

1.2 Augmented Reality 3

1.2.1 Types of AR display.....5

1.3 Vehicle head-up displays.....5

1.3.1 Vehicle HUDs compared to traditional head down displays 8

1.4 Structure of the visual system 11

1.4.1 Central and peripheral vision.....13

1.4.2 Useful field of view. 15

1.5 Effects of presenting HUDs over a wider field of view 19

1.6 Human visual attention 20

1.6.1 Early vs late selection processing theories 22

1.6.2 Perceptual load and its impact on information processing.....28

1.6.3 Cognitive control and perceptual load of attetnion 31

1.7 Alterniative approaches to peripheral visual field processing.....35

1.7.1 Cognitve tunneling vs general interference 37

1.7.2 Tunneling compared to load theory.....38

1.7.3 What is the effect of display area? 40

1.7.3.1 Driving simulator research in UFOV 41

1.7.3.2 Large displays without driving related task.....45

1.8 Perceptual issues of depth perception in AR 48

1.9 How do humans judge depth 50
1.9.1 Stereopsis: depth information derived from two eyes.....50
1.9.2 Optical Cue Integration 51
1.9.2.1 *Reliability and optimal cue combination* 51
1.9.2.2 *Reliability and viewing geometry* 53
1.9.2.3 *Measuring depth resolution and obtaining PSE* 56
1.9.2.4 *Signal in noise & ideal observer* 57
1.9.2.1 *Models of optimal cue combinaiton* 59
1.10 Human intermediate distance perception in the real-world. 62
1.10.1 Binocular disparity at intermediate distance testing. 63
1.11 How does depth perception differ in augmented enviroments?. 66
1.11.1 How depth has been measured in VR and AR. 70
1.11.2 The use of HMDs 76
1.12 Summary. 77
Chapter 2: Field of view and attentional workload. 80
Section A: Workload and horizontal eccentricities 82
2.1 Introduction.....82
2.1.1 Attetnional Tunnelling.....83
2.1.2 Perceptual load theory predicitions 83
2.1.3 Research Question & Hypothesis.....85
2.2 Method.....86
2.2.1 Participants 86
2.2.2 Stimulus & Apparatus 87
2.2.3 Design.....89

2.2.4 Procedure 91
2.3 Results.....92
2.3.1 The impact of foveal load on peripheral perfomance 92
2.3.1.1 *Peripheral discrimination perfromance* 92
2.3.1.2 *Reaction times for peripheral targets*.....98
2.3.2 Critical Tracking Task 100
2.4 Discussion.....102
2.4.1 Cognitve load increases peripheal perfromance decreases 102
2.4.2 Left visual field deficit 104
2.4.3 Protection of CTT ? 106
2.4.4 Study limitations..... 107
2.4.5 Summary & Conclusion 108
Section B: Increased presentation time and vertical eccentricities 109
2.5 Introduction.....109
2.5.1 Vertical target presentations 109
2.5.2 Peripheral target presentation time 110
2.5.3 Research Question & hypothesis.....111
2.6 Method.....112
2.6.1 Participants 112
2.6.2 Design 112
2.6.3 Procedure 114
2.7 Results.....114
2.7.1 Increased foveal load on peripheral task performance 114
2.7.2 Increased peripheral load on foveal task performance 120

2.7.3 Reaction Times 120
2.7.4 Peripheral target appearance 122
2.7.5 Section A and Section B comparisons 123
2.8 Discussion 125
2.8.1 Early or late attetnional selection 125
2.8.2 Fixation task priority 126
2.8.3 Horizontal and vertical eccentricities comparison 127
2.8.4 Implications for AR design with large display size 128
2.8.5 Summary and conclusion 128
Chapter 3 : Depth perception in augmented reality 130
Section A : Stereopsis AR depth perception 132
3.1 Introduction 132
3.1.1 Comparative depth judgments in AR, VR and real-world 132
3.1.2 Vertical image location in the visual field 134
3.1.3 Horizontal image location in the visual field 136
3.1.4 Overlaid virtual imagery 136
3.1.5 Summary and Rationale 137
3.1.6 Research Question & hypothesis 138
3.2 Method 139
3.2.1 Participants 139
3.2.2 Experimental Design 139
3.2.3 Showing the AR image 140
3.2.5 Image height 141
3.2.6 Procedure 142

3.3 Results.....143
3.3.1 Perceptual equality of AR depth perception.....147
3.3.2 Virtual image percieved in front.....148
3.3.3 Virtual image percieved behind.....149
3.3.4 Slope analysis 150
3.3.5 Reaction Times 152
3.4 Disccusion.....153
3.4.1 Depth error in AR 153
3.4.2 Underestimation in AR.....154
3.4.3 Virtual image position 155
3.4.4 Individual difference & slow response times 156
3.4.5 Summary & Conclusion 156
Section B : Relative cues of size and brightness in AR depth perception 158
3.5 Introduction.....158
3.5.1 The importance of size and brighntess depth cues in virtual and augmented enviroments 158
3.5.2 Summary & Rationale 162
3.5.3 Question & Hypothesis.....163
3.6 Method.....164
3.6.1 Participants 164
3.6.1.1 *Size*.....164
3.6.1.2 *Brightness* 164
3.6.1.3 *Size and brightness combined* 164
3.6.2 Design 164

3.6.3 Virtual image manipulations 165
3.6.4 Procedure 165
3.7 Results.....166
3.7.1 Which depth cue produced the best depth estimate 166
3.7.2 Experiment 1 : Size 167
3.7.3 Experiment 2 : Brightness 168
3.7.4 Experiment 3 : Size & brightness.....169
3.7.5 Slope analysis 170
3.7.5.1 *Size*.....170
3.7.5.3 *Brightness* 171
3.7.5.3 *Size and Brightness*.....171
3.7.6 Reaction times 171
3.8 Discussion.....173
3.8.1 Brightness not signifcantly infeior to size as a depth cue 174
3.8.2 How results are understood in terms of cue combination 174
3.8.3 Overlaid presentation produce less error than above 175
3.8.4 Reaction times 175
3.8.5 Study limitations.....176
3.8.6 Implications for AR devices 176
3.8.7 Summary & conclsuion 177
Section C : How is depth information acquired in AR : the virtual image or the real-world background ?.....178
3.9 Introduction.....178
3.9.1 Virtual image or real-world background.....178

3.9.2 Research Question & hypothesis 180
3.10 Method 181
3.10.1 Participants 181
3.10.2 Experimental design 181
3.10.3 Procedure 182
3.11 Results 183
3.11.1 Real-world background or virtual image cues 183
3.11.2 Slope Analysis 188
3.11.3 Reaction times 188
3.12 Discussion 189
3.12.1 Depth information gained from either the surrounding world or the virtual image 190
3.12.2 Summay & conclusion 191
Chapter 4 193
4.1 Introduction 193
4.2 Wide field of view and attention 193
4.2.1 FOV and attetnion conclusions in context of perceptual load theory 194
4.2.2 FOV and attetnion findings for design of AR devices 197
4.2.2.1 *Peripheral accuracy* 197
4.2.2.2 *Left and right visual field* 197
4.3 Depth perception 199
4.3.1 How does depth perception in AR compare to real-world 200
4.4 Conclusion 201

4.4.1 Contribution to knowlwedge 201
4.4.1.1 *Attention and wide FOV* 201
4.4.1.2 *Depth perception in AR* 202
4.4.2 Limitations of studies 202
4.4.3 Future directions 202
4.4.3.1 *Switching attenion between depth planes* 202
4.4.3.2 *Recommendations for practice (vehicle manufactures)* 204
4.4.3.2.1*Safety* 205
4.4.3.2.2 *Normal and abnormal conditions* 206
4.4.3.2.3 *Reliability & Quality* 206

Abstract

There has been a recent development and growth of augmented reality devices across the world. The technology has the potential to enhance a user's perception and experience of the environment which surrounds them. For designers and engineers of 3D displays (e.g. vehicle HUDs) the technology offers an opportunity to form a more enriched and exciting experience for drivers. A key anticipation of AR is the capability to present virtual information at depth, as well as on displays with a wide field of view. Applying psychological theories and models of visual attention (e.g. perceptual load theory) and depth perception (e.g. modified weak fusion), the aim of the study was to establish human abilities with AR advancements in depth and useful field of view (UFOV). Conclusions were applied to the development of psychological theory and helpful design guidelines for AR designers and engineers. This has been done by creating two sets of experimental paradigms. The first set of experiments presented peripheral Landolt C target stimuli on a large display (120^{o}), simultaneously whilst participants controlled a central tracking task (CTT). Over two experiments workload was manipulated on both central and peripheral tasks, including presentation time, contrast, tracking speed, and size. Results indicated that peripheral performance did not decrease as a function eccentricity. Conclusions suggested a demonstration of attentional selection dependent on stimulus parameters, with performance potentially dependent on more temporal characteristics than eccentricity. 'Best-practice' design guidelines are presented for AR HUDs with large display size. The second design paradigm implemented a two-alternative forced-choice psychophysical depth judgment task, assessing thresholds for correctly determining the depth of a virtual diamond in reference to a real-world object. Over a series of experiments a number of parameters were assessed including virtual image height in the visual field, virtual image cue properties (relative size and brightness), and background scene manipulations. Depth thresholds for the AR image were tested with the pedestrian target at 5m, 10m, 20m and 25m. Results demonstrated that depth resolution is poor in augmented reality compared to real-world environments, with conclusions indicating a potential difference in how relative cues of depth interact with each other in AR environments, i.e. in AR cues from virtual imagery and real-world background may not be effectively combined together. Design guidelines regarding where in depth a virtual image needs to be placed to 'match' that of a specific real-world object, and how relative cues to influence depth perception are presented.

Glossary & Abbreviations

Augmented Reality (AR): Technology system where computer generated graphics and other digital information are superimposed onto a real world environment

Critical Tracking Task (CTT): a lab based task which has been used as a measure of driving performance (i.e. in-lane keeping). The task involves participants controlling a white virtual bar within the boundaries of a larger moving black bar. The task is to keep the controlling white bar within the boundaries of the larger black bar

Field of View (FOV): the range of the observable world that is perceived at any point

Head-up display (HUD): A transparent display which presents virtual visual information directly in front of the user; developed by aviation technology and has been manufactured in road vehicles

Head-down display (HDD): Displays which require users to look away from point of fixation. For example and instrument panel or dashboard in a road vehicle

Head-mounted display (HMD): Display which is worn on a user's head such as a helmet or a pair of googles and can be used to present AR

Useful Field of View (UFOV): The spatial range over which information can be effectively combined

Chapter 1: Introduction

1.1 Research Motivation

The current research was co-funded by the University of Nottingham (School of Psychology), EPSRC, and the automobile company Jaguar Land Rover (automotive human machine interface). The aims of the project were two fold. One aim was to gain psychological understanding of visual attention, field of view and depth perception for augmented reality devices. The second aim was to produce some design guidelines for future consideration in developing vehicle AR HUDs.

For this reason experiments presented in the thesis were designed in such a way to assess both ambitions of the research institutions and vehicle manufacture by applying psychophysical measurement techniques to produce Human Factors design guidelines for vehicles. The design of the experiments in the thesis importantly permitted for psychological theories of attentional workload (within the framework of field of view) and models of optimal cue combination (in depth perception) to be tested and developed for new visual technologies, specifically augmented reality devices.

1.1.1 Jaguar Land Rover

For a number of years vehicle manufactures have been developing techniques to potentially present 3D visual information across potentially an entire vehicle windscreen and presented at a specific depth in the environment (AR HUDs). It is essential for vehicle manufactures to produce guidelines (do's and do not's) for the design of 3D displays to prohibit more innovative HMI designs (Pitts, Hasedzic, Skrypchuk Attridge & Williams (2015). With an increasing consumption of information throughout worldwide populations,

drivers do expect superior user experiences and visuals whilst driving (Bostrom & Ramstrom, 2014).

Vehicle designers will need to understand how users respond to technologies which command a large amount of visual space; both in terms of field of view and depth. This will help determine a 'design space' where information can be projected? Where in depth does a virtual image need to be placed to match a corresponding real-world object it is highlighting? Not only will designers need to create technologies that meet expectations visually, they will need to create a system which is safe and not distracting for drivers.

1.1.2 Theoretical motivation

Examining how users may process visual information on a device commanding a large volume of a user's visual space could help further understand the nature of selective processing of irrelevant information. Users of prospective AR HUDs will have to exclude potential distractors whilst selecting relevant information from across their FOV.

There has been a long held debate in attentional theory about the specific processing stage that attentional selection becomes active (see 1.6.1); referred to as early and late selection. The thesis assesses the attentional debate and the impact of perceptual load (Lavie, 1995, 2005, 2010), by examining how variables such as position in space (fixed eccentricity) and time duration of irrelevant stimuli affects human ability in selective processing of irrelevant information.

The capability to present visual information at depth may allow for faster registration to highlighted objects in a scene. However previous research has found that the most common perceptual difficulty in an AR environment is a poor interpretation of depth (Kruijff, Swan II & Feiner, 2010). By examining a variety of cues to depth in an AR environment, it can be understand why in

certain environments human depth perception is limited. There is a potential to produce models of depth perception which extend current theories. In turn this will aid designers in the cues to depth to utilise for future devices.

1.2 Augmented Reality

Augmented Reality (AR) superimposes computer-generated graphics, and other digital information onto the real-world environment; opening up a whole new experience for users by creating a more enriched visual environment.

Over the last few years AR has been applied to handheld devices (smartphones), wearable computers (head-mounted displays, google glasses) and large in-built devices (head-up displays); for applications as varied as health (surgery), education, marketing, telecommunications, military applications, vehicle safety, and entertainment (Yeh & Wickens, 2000; Haller, et al, 2006). Examples include laparoscopic surgery (e.g. keyhole surgery), tablet and smartphone apps for the classroom and entertainment (‘Aurasma,’’ZooBurst’’), novel personalised advertising for consumers, and vehicle head-up displays (Pisano et al, 1998; Dunleavy, Dede & Mitchell, 2009*;* Gish & Staplin, 1995). Not only has the popularity of AR increased, but so has the advancement in technology. Over the coming years AR will be able to present visual information on larger devices allowing for access of a wider field of view (FOV) for the user, and presenting virtual imagery at increasing and multiple depths; an increase in both two-dimensional and three-dimensional perceptual space.

With this new technological advancement perceptual difficulties have arisen; specifically with regards to visual attention and depth perception. A number of research studies have demonstrated issues regarding inefficient uses of attention (Weintraub, 1992; Gish & Staplin, 1995)**.** As AR enables virtual imagery to be overlaid on to real-world environments, problems with matching

the virtual image and target object in depth have been documented; resulting in accommodation and convergence issues and difficulties in accurately conveying depth (Poitscke, Ablassmeier & Rigoll, 2008; Jones , Swan II, Singh, Kolstad & Ellis, 2008; Drascic & Milgram, 1996).

Also in providing additional information to the user AR, as well as being a potential aid to guide attention, overlaid virtual imagery can obstruct a user. For example vehicle head-up displays (HUDs) can present text messages to highlight points of interest in the driving scene (e.g. a street name, building name, directions). When a user focuses their attention on to the text stimuli, the interpretation of the semantic meaning of the words will need to be understood, thus potentially ignoring other vital stimuli in the driving scene.

The thesis explores issues of visual attention and depth perception in AR. Chapter 1 of the thesis is split into two sections; the first concentrating on visual attention and FOV with regards to AR, the second depth perception in AR.

The first section will begin by describing AR applications; specifically vehicle head-up displays, followed by detailing the structure of the visual system and attentional networks within the brain. Following on, concepts of attentional load and dual task are explained, and the debate between early and late attentional selection is discussed. The section finishes by examining how the theories of human attention have been explored in the context of AR research. The second section of the chapter opens by focussing on the issues raised for depth perception research in AR, then will describe how humans perceive depth, detailing distance perception research in real-world environments. Finally the chapter will finish by examining research findings for depth perception in AR.

1.2.1 Types of AR display

There are two different variations of visual AR displays; optical-based and video-based (Barfield & Caudell, 2001). Optical-based displays involve lenses which permit light to pass through the user's eyes allowing the actual object to be perceived as normal (see *Figure 1.1*). These lenses can superimpose computer generated imagery on the real-world scene through reflection onto the lens from a small display. As a negative point images cannot fully occlude the real world image and appear more as a 'watermark' (Jerome, 2007).

The alternative video-based AR displays are perceptually very different as the light from the objects is not sent to the eyes of the user. Instead real-world images are captured by cameras and reconstructed on small display presented in front of the user (see *Figure 1.2*). In essence the computer blends computer generated images on to the real-world environment. The benefits of these displays over optical is that the images can be seen clearly and fully occluded real-world objects, however just as in a computer monitor the display refreshes causing potential eye strain, headaches and simulator sickness (Jerome 2007).

One of the aims of the thesis is to apply experimental findings to vehicle AR HUDs, which is type of optical see-through display. The working of a HUD starts with a light source from the picture generation unit (see *Figure 1.3*), which hits a fold mirror, and then reflects on to a rotatable mirror which projects onto a windshield via a glare trap.

1.3 Vehicle head-up displays

An example of the exploration and possibilities of applying AR is in the use of vehicle Head-up Displays (HUDs). HUDs are a system which projects driver information on the windscreen directly in front of the driver, increasing the amount of time a driver's eyes are spent on the road; resulting in faster

reaction times to both hazards and warnings (Charissis, Papanastasiou & Vlachos, 2008; Liu & Wen, 2004).

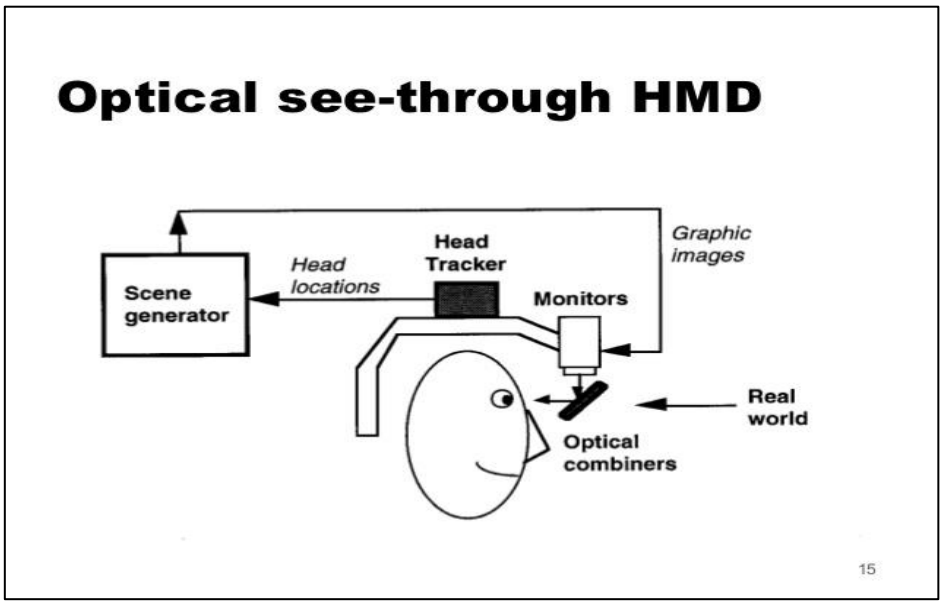


Figure 1.1: Schematic presentation of optical see-through AR display https://www.slideshare.net/Khyati14Ganatra/augmented-reality-ppt-47315337

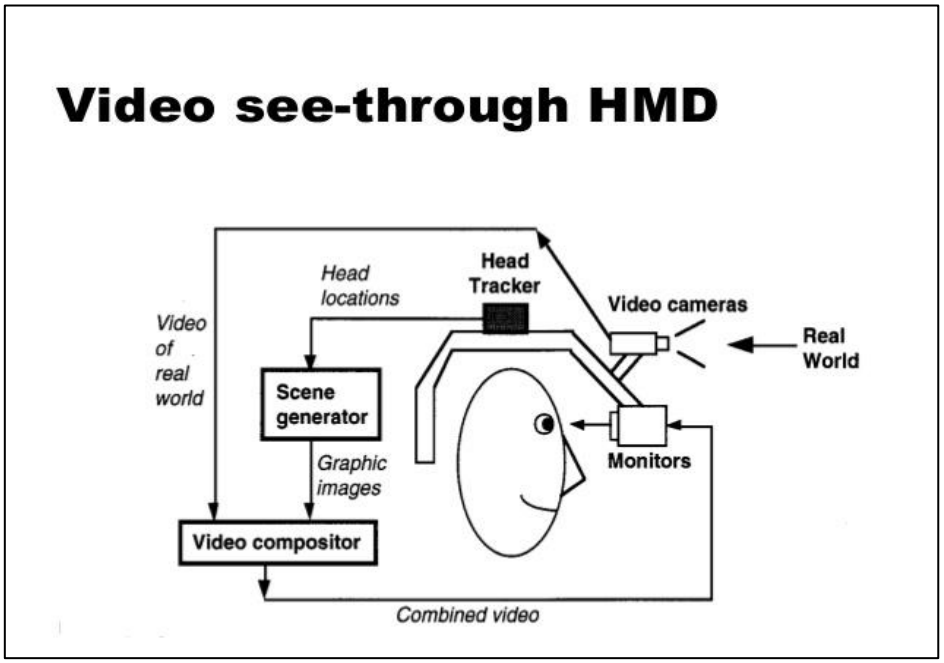


Figure1.2: Schematic presentation of video see-through AR display https://www.slideshare.net/Khyati14Ganatra/augmented-reality-ppt-47315337

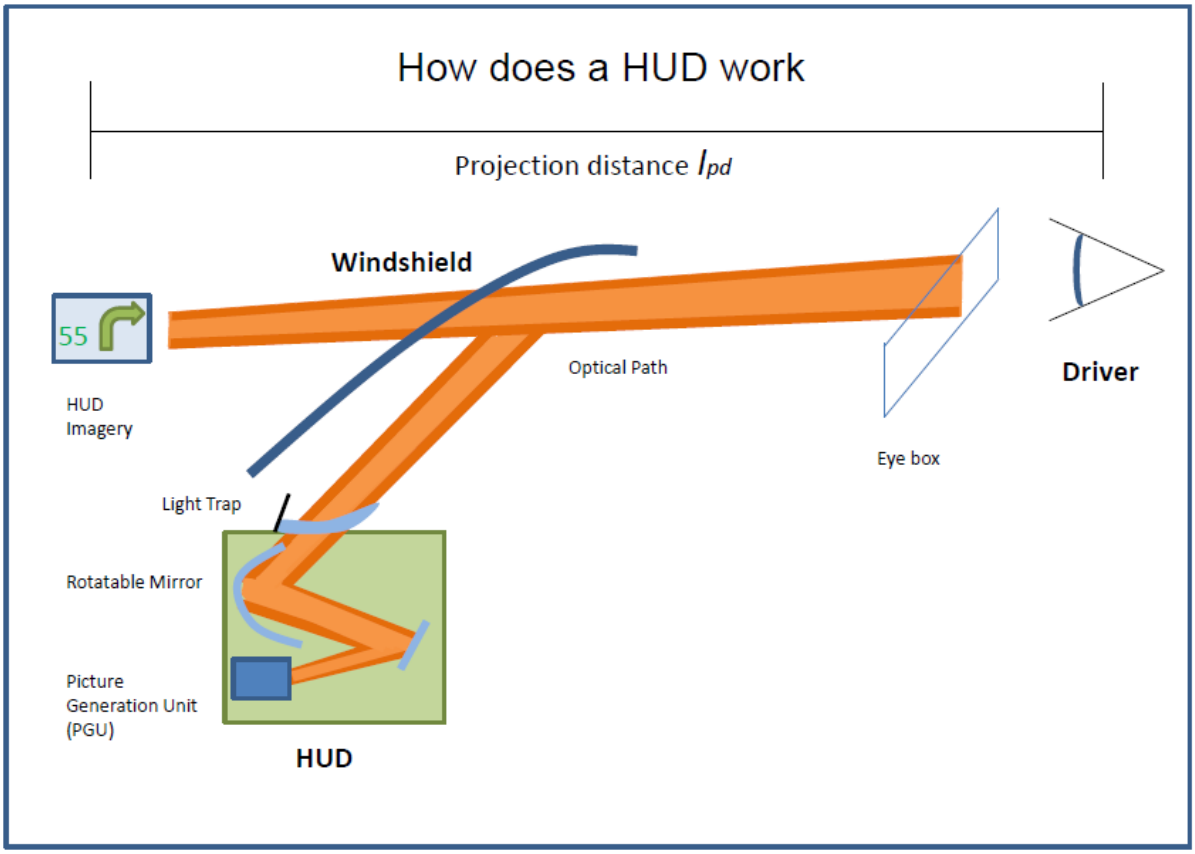


Figure 1.3: Schematic of how a vehicle HUD works.

Driving has been typically considered as a combination of three subtasks: vehicle control, hazard detection, and navigation (Dewar, Olson & Alexander, 2002; Wickens, Gordon, Liu & Lee, 1998). For the past few decades car manufactures have been developing in-vehicle technologies to best present visual information to drivers; with the aim to produce improved levels of driver performance (vehicle control, hazard detection and navigation). Conventionally, in-vehicle visual information has been presented on instrument panels and clusters (Head-down displays, HDDs) which require the drivers to take their eyes-off-the-road. Development has worked on displays which present visual information directly in front of the driver on to the windscreen. Initially developed from aviation technology, from gun-sights on military aircraft to radars on commercial planes (Weintraub, 1992), HUDs project visual information onto a position directly in front of a driver's field of view (FOV) on to an automobile windscreen. With all the vital driving information presented

right in front of the driver, HUDs aim to significantly reduce the amount of time a driver's eyes are off the road compared to traditional Head-Down Displays (HDD); resulting in better and faster detection of objects & events in the driver's external scene (Tufano, 1997).

Vehicle HUDs are an example of highlighting the possibilities of what AR can do; with the large windshield allowing for a large FOV and the potential for highlighting numerous types of information at multiple depths. The thesis will explore how psychological models and theories of visual attention and depth perception can help in the advancement of AR applications. As well what can psychology learn and understand about our own knowledge when viewing a world which combines the real and the virtual. This will be done by testing AR-like and HUD-like systems in in vehicle like settings.

1.3.1 Vehicle HUDs compared to traditional head down displays

Several studies have found that HUDs have advantages over HDDs. For instance they reduce visual reaccommodation between viewing the instrument and the outside world, reduce the amount of visual scanning required, give superior navigating performance, speedier responses to in-vehicle warnings and to urgent external events, and give general improved vehicle control, less mental stress and higher user satisfaction (Liu & Wen, 2004; Burnett, 2003; Weinberg, Harsham & Medenica, 2011)**.**

For example Liu & Wen (2004) compared HUDs and HDDs for a commercial lorry driving simulator experiment. Comparing driver reaction times, the study found significantly faster responses for driver using HUDs for tasks requiring navigation and response to urgent events. The experiment also found speed control (measure by variation of speed) was more consistent in HUD drives compared to HDDs.

Weinberg et al (2011) conducted a simulator experiment involving a street name finding task. Again comparing HUDs against HDDs, the study found twice as many off-road glances for users of HDD compared to HUDs. Analysing participant feedback revealed the HUD scored higher in user satisfaction and a lower impact of mental load.

In another driving simulator experiment Burnett (2003) found evidence for improved navigation for HUDs compared to HDDs (fewer errors), although did conclude that HUDs may produce some negative outcomes; such as awkward driving postures to read HUD imagery. Also there were concerns from participants about the constant presence of information presented to them (possibly distracting).

After reviewing early experimental research on HUDs, Gish & Staplin (1995) raised concerns about the use of automobile HUDs; with potential issues regarding workload, visual clutter (obscuration of the road scene) and depth focussing (Gish & Staplin, 1995).

Early research suggested that that the advantages HUDs brought were only associated with simple displays, under moderate ambient light levels and conditions of low workload (Okabayashi, Furukawa, Sakata and Hatada, 1992; Weintraub and Ensing, 1992; Larish & Wickens, 1991). It was also suggested that HUDs may already add to a drivers demanding workload of vehicle control (Gish & Staplin, 1995). As HUDs were limited in FOV concerns about HUD imagery obscuring a driver's vision of the external scene, and interfere with aspects of the primary task of driving were raised (Okabayashi, et al, 1992).

Further concerns have been linked to the depth focussing required for a HUD image. Vehicle HUDs require that drivers' eyes are focussed at distances much closer than the real-world objects highlighted. Researchers were cautious of a 'misconvergence trap;' double vision (diplopia) of faraway objects

(Tufano, 1997). The proposed focal distance of a HUD was 2m. Based upon critical reviewing of all work in aviation and automobile HUD, Tufano (1997) expressed concerns about the distance selected for car HUD imagery; with drivers having possible vergence difficulties. This could result in micropsia; where objects appear smaller than they really are. Sojourner & Antin (1990) made early advisory research areas which need to be addressed for vehicle HUDs; the impact they have on perceived size and distance of real-world objects in the external scene.

Currently, HUD devices have tended to contain basic driver information, previously presented on the head down display, including speedometer, RPM, fuel levels and vehicles warnings. This information is presented within a small field of view (around 20cm x 30cm), single colour representations and fixed driver image depths of around 2 -2.5m. To solve many of the early concerns mentioned above (incorrect depth focussing, workload and visual clutter), vehicle HUDs may increasingly present virtual images that are overlaid onto real world scenes (at the perceived depth of the object). These images can be designed to provide navigation information or collision warnings (Plavsic, Bubb, Duschl, Tonnis & Klinker, 2010; Lim, Newstrom, Shankwitz & Donath, 2000).

Augmented reality will be implemented into devices capable of large spatial perception; both in terms of width and length (for instance vehicle windscreens). There have been concerns raised in AR perception about correctly presenting depth (Swan II, Jones, Kolstad, Livingston & Smallman, 2007), and attentional workload issues when implanting AR devices to aid in already complex tasks; such as driving (Gish & Staplin, 1995**).** The thesis will present current theories of attentional workload (see 1.6.2), modelling of depth cues in the real-world (see 1.9.2), and how perception in augmented reality can add to our current knowledge.

A further development for future HUDs in vehicles will be the field of view. Augmented reality HUDs will be implemented into devices capable of projection across wide spaces, in terms of width and length (for instance across the whole vehicle windscreen). Furthermore, AR images may be presented over a wider range of depths. These developments present two potential issues.

First additional of multiple displays to HUD may raise attentional workload issues when implanting AR devices to aid in already complex tasks; such as driving (Gish & Staplin, 1995). Secondly, there have been concerns raised in research around AR about correctly presenting depth (Swan II et al, 2007), there are gaps in the current literature around how AR depth information is integrated with the real world; how do depth cues interact with each other in AR at far distances and is virtual and real-world combined effectively. The thesis will present current theories of vision and attentional workload (see 1.5), and modelling of depth cues in the real-world (see 1.10).

1.4 Structure of the visual system

Prior to the discussion of visual attention and depth perception, it is necessary to briefly review the structure of the visual system and perceptual abilities in different parts of the visual field.

The human visual system comprises of the eyes (sensory organs), which detect light rays which are then coded into visual information; before being transmitted to the brain for analysis. Light rays enter the eye, pass through the cornea, anterior chamber, lens, and vitreous and finally focus onto the retina (containing photoreceptors) where visual processing begins (Polyak, 1941). The photoreceptors contained within the retina are activated by light (Hecht, Shlar & Pirenne, 1942; Baylor, Lamb & Yau, 1979). The photoreceptors modulate movement of bipolar cells; which synapse with ganglion cells forming the optic

nerve. Information is carried to the brain through the optic nerve (Neveu & Jeffery, 2007).

The optic nerve continues eventually meeting and crossing over at the optic chiasm (see *Figure 1.4*). The axons from the retina cross over such that information from the left eye travels through the right optic tract, with information from the right eye travelling via the left optic tract (Mcllwain, 1964). Both optic tract pathways meet and terminate at the lateral geniculate nucleus (LGN); the visual part of the dorsal thalamus. The LGN is the primary relay nucleus for visual processing and is first stage of coordinating vision to the visual cortex from the eyes (Derrington, Krauskopf & Lennie, 1984). The visual cortex maintains the final processing of neural signals from the retina.

The visual cortex consists of six separate regions; V1, V2, V3, V3a, V4, and V5 (Watson et al, 1993). V1, is often referred to as the primary visual cortex, and is the first region in visual information processing; the area where neurons from the LGN synapse. The role of the primary visual cortex (striate cortex) is to project information to other areas of the cerebral cortex (extrastriate cortex) which is involved in more complex visual perception; visual processing occurs in a hierarchal process (Glickstein, Rizzolatti, & Gennari, 1984). V1 has a well-defined spatial map for visual information, processing information primarily from the fovea (cortical magnification); signals containing information on orientation of objects, colour, and form (Glickstein, et al 1984). After V1 neural signals move through to V2 for further colour and form processing (Shipp & Zeki, 1985). Continuing through the visual cortex processing becomes more associative with signals for motion of objects, self-motion and spatial reasoning involved in later regions of the visual cortex (Watson et al, 1993). V5, or the temporal visual areas, is where motion is processed (Watson et al, 1993).

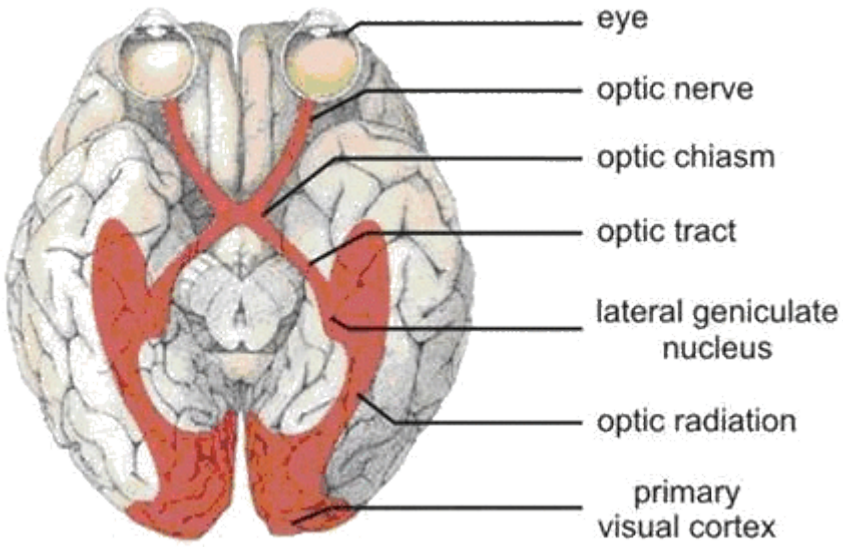


Figure 1.4: An illustration of the cross section of the human visual section (https://www.jjkvc.org/visual-system)

1.4.1 Central and peripheral vision

The finest visual resolution (quality of image) is perceived through the very narrow fovea (central vision) of the retina (Millodot, 1972). Central vision consists of the fovea (5.2^{o}), parafovea (5^{o}-9^{o}) and perifovea (9^{o}–17^{o}); these three subsections form the macula (Osterberg, 1935), or paracentral vision. The further away information is presented away from the center of our gaze, the more blurred or distorted the image becomes (See *Figure 1.5*). Peripheral vision can be regarded as outside paracentral vision (18^{o}>); which is split into near peripheral (18^{o} - 30^{o}), mid peripheral (30^{o} - 60^{o}) and far peripheral (60^{o} - 180^{o}) (Besharse and Bok, 2011; Abramov, Gordon and Chan 1991; Strasburger, Rentschler & Juttner, 2011).

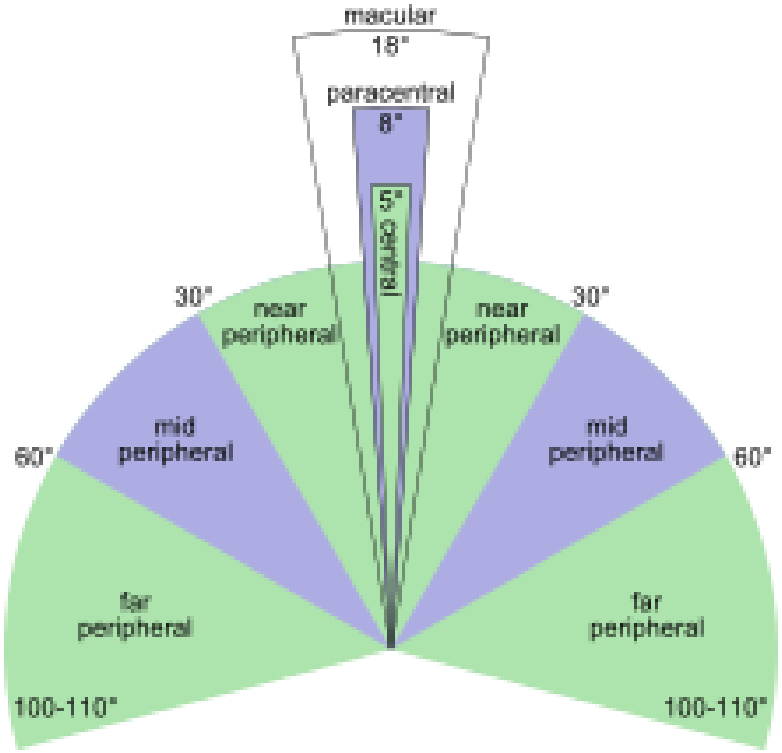


Figure 1.5: Diagram illustrating the spatial range of the human visual field (https://en.wikipedia.org/wiki/Peripheral_vision#/media/File:Peripheral_vision.svg

Physiologically central and peripheral vision is very different. Psychophysical research has examined the response properties and the sensitivity for stimuli with fixed spatial properties (size, spatial frequency, contrast) decreases as eccentricity increases (Anstis, 1974; Rovamo, Virsu & Naesaenen, 1978; Virsu and Rovamo, 1979). The large concentration of rods and cones found in the fovea decreases in the peripheral regions of the retina, which is responsible for the poor resolution and degradation in visual acuity, contrast sensitivity, colour, shape and text vision in peripheral vision (Daitch & Green, 1969; Strasburger, Rentschler & Juttner, 2011). In the primary visual cortex a larger number of cells respond to information in central vision compared to peripheral vision; referred to as cortical magnification (Daniel and Whitteridge, 1961). Cortical magnification has been found to vary across a

range of factors including target stimulus (e.g. Landolt C vs Gabor patches), target position and individual differences (Strasburger et al, 2011).

These fundamental structural and perceptual differences between central and peripheral vision have real-world implications; tasks requiring discrimination of fine detailed information are processed by central vision (Henderson & Hollingworth, 1999; Nuthmann, 2014). Peripheral vision guides attention to fixated objects whilst identifying and discriminating peripheral objects (Deubel & Schnider, 1996; Henderson 1992; Thorpe, Gegenfurtner, Fabre-Thorpe & Bulthoff, 2001). It is also understood that relevant information required for driving is obtained from both central and peripheral vision (Obayakashi et al, 1991; Miura, 1990).

1.4.2 Useful field of view

The spatial range over which information can be effectively combined is commonly referred to as useful field of view (UFOV). The concept was first proposed by Sanders (1970), whom applied the term 'functional visual field;' the size (area) of visual field where extraction of visual information can be taken from a single glance (Ball, Wadley & Edwards, 2002).

Sanders (1970) presented participants with a simple four-choice discrimination task. Participants were presented six columns of dots spread along a horizontal row. Each column contained four dots; except one single column, which totalled five dots. The participants' task was to determine which column contained five dots. A new set of columns appeared directly after the participant responded. The column of dots was presented at three display angles (68^{o}, 81^{o} and 94^{o}). Measuring response accuracy and reaction times, results showed a decrease of performance as a function of display angle. The increasingly dispersed stimuli required more complex processing; additional processes required to accurately judge stimuli. Sanders (1970) divided UFOV

into three areas; stationary field, eye field and head field. Stationary field uses near peripheral vision (20^o - 40^o) to perceive a stimulus. When eye movements are necessary for target detection, it is in the eye field (40^o - 90^o). Head field refers to situations where a target falls beyond useful eye movements and requires the head to reposition to detect the target ($<90^o$).

To assess the area and function of UFOV a computer-based task was developed; the UFOV assessment (Ball, Beard, Roenker, Miller & Griggs, 1988). The test includes three subsets; performance on a central visual identification task, ability to divide attention between central and peripheral stimuli, and to ability to select peripheral stimuli from amongst a set of distractors. The test first requires participants to identify a target presented in the center fixation point of the screen. This is followed by identification of center target and simultaneously a target presented in the periphery of the display. In the final part of the experiment participants must identify in the periphery target objects which are presented amongst distractors (see *Figure 1.6*).

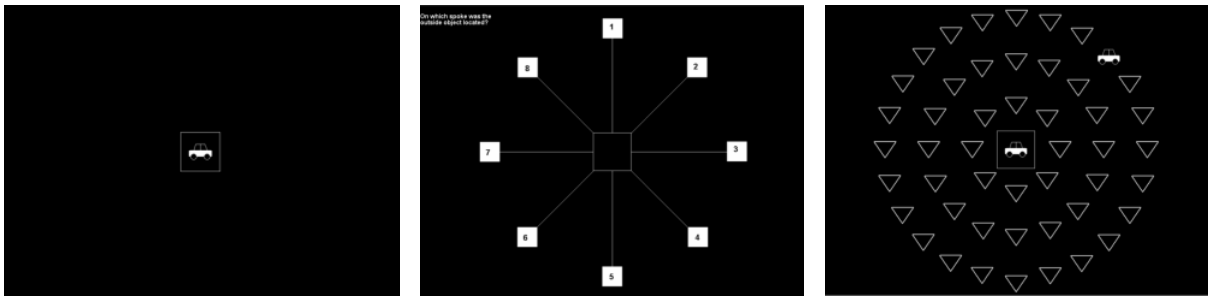


Figure 1.6: Schematic diagram for the procedure of the FOV assessment (Ball et al, 1988)

The display area FOV was 60^o. Central fixation presentation time was 750msec, followed by a radial pattern with eight numbers (see *Figure 1.6).* The pattern remained visible until the participant responded.

UFOV in the stationary field has been used as an assessment of parallel attentional processing, and has been claimed to correlate with mental status and measure of cognitive ability (Edwards, Wadley, Myers, Roenker, Cissell, & Ball, 2002). For instance Edwards et al (2002) examined cognitive influences on UFOV; this was done by administering a UFOV assessment and a variety of questionnaires and visual attention assessments; 'road sign test,' 'digit symbol substitution' a self-rated health system.

A digital symbol task (Wechsler, 1981; Tun, Wingfield & Lindfield, 1997) is an assessment of psychomotor speed (actions demonstrating fine motor skills). The task involved presentations of a key which is paired to digits ranging from 1-9 with symbols. The participants' task was to view a series of digits and draw each digit which matches the appropriate symbol. Participants had 90seconds to complete as many substitutions as possible.

The 'road sign test' was a lab based assessment which examined everyday visual search skills (Ball & Owsley, 2000). Participants were presented on a computer screen road signs (pedestrian, bicycle, right and left turn arrows), with a red slash through (distractors) or without (targets). Participants were instructed to react to targets and ignore distractors. When presented target left and right turns signs, participants had to move a mouse in the direction indicated by the sign. For pedestrian and bicycle signs participants were required to click a mouse button. Reaction times for correct responses were recorded. Participants also were required to self-report their health status with regards to heart disease, hypertension, depression, arthritis, cancer, stroke, as well as any eye conditions; glaucoma, macular degeneration, cataracts, dry eye syndrome.

Analysis found that UFOV performance was significantly influenced by factors such as age, mental status and vision.

Also it has also been claimed by researchers that UFOV performance is dependent on the reliability of visual sensory information and an individual's higher order processing abilities (Ball et al, 1990; Owsley, Ball, & Keeton, 1995). Owsley, Ball and Keeton (1995) carried out UFOV assessment on older adults. Results of the experiment demonstrated that participants have difficulty locating peripheral objects despite having good acuity and contrast sensitivity. Owsley et al (1995) concluded impairment of higher order visual processing mechanisms is the major cause of visual search issues.

Applying knowledge of the human visual system, investigations have tested the impact of reduced UFOV on observers; showing that as UFOV decreases in area there is a correlation in the increase in errors and accidents (Owsley et al, 1998; Owsley, Ball, Sloane, Roenker & Bruni 1991).

Sanders further proposed that the area of a person's UFOV is a function of four variables; target presentation duration, competition between central and peripheral tasks, the salience of a peripheral target and also the eccentricity of distance of the peripheral target from central vision (Sanders, 1970). UFOV is accepted as a crucial component in perceiving relevant objects whilst driving (Hills, 1980), therefore creating vehicle displays which increase UFOV for drivers could be seen as potentially beneficially for a driver.

As discussed above, vehicle HUDs are likely to present AR images over a larger space in the near future. Users may be required to split attention between central and peripheral vision to in order to maintain safe vehicle control (Gabbard, Fitch & Kim, 2014). In this thesis a set of experiments is presented (see chapter 2) examining peripheral target presentation duration of peripheral targets at large eccentricities, whilst completing a central task. In the next section I review previous research on presenting images over a large area.

1.5 Effects of presenting HUDs over a wider Field of view

Steinfeld (1995) was one of the first researchers to experiment on the potential of a full-windscreen HUD. In a mock-up car setup (based on experimental setup by Green & Williams, 1992) the researchers compared reaction times to information displayed on an instrument panel (HDD) and a full windscreen HUD. Participants were presented slides of residential intersections. Simultaneously a slide of a navigation display was presented either on an instrument panel or superimposed onto the scene to simulate a full windscreen HUD. The participant's task was to indicate if the two images were the same or different type of intersection. The study found results of faster response times for the full windshield HUD compared to HDDs, for the intersection images presented.

Steinfeld (1995) compared there results to a similar previous study which compared small HUDs to a HDD (Williams & Green, 1992). The comparison demonstrated faster responses of the full windscreen HUD compared to the small HUD. However when comparing rates of error, full windshield HUDs produced twice as many incorrect responses as compared to small HUDs (Steinfeld, 1995).

Participants' subjective preferences indicated that a full windshield HUD was the most favoured display. The results of this study indicated that there might be improvements in performance for HUDs which present their images over the full windscreen. The conclusions from this study are limited, however, by the lack of a direct comparison between wide and narrow HUD displays and the lack of a specific driving related task.

Charrisis, Papanastasiou and Anderson (2010) in a driving simulator compared a full windshield simulated HUD prototype with contemporary HDDs by testing drivers on a variety of driving scenarios, with both low and high

difficulty scenarios. Task difficulty was manipulated by traffic density. When comparing reaction times and the occurrence of driving errors (e.g. going over the speed limit, crossing lane position), analysis indicated that the HUD interface reduced response times and driver errors compared to HDD.

Simulation testing full windshield HUDs have claimed the device provides the driver with a greater amount of useful driver information, potentially adding to driver performance, especially with regards to overall driving performance and the chances of accident avoidance (Charrisis, Naef, Papanastasioum, Patera, 2007; Charrisis, Naef, Arafat and Vlachos, 2008). However these conclusions have been made without direct comparisons with other devices. For instance Charrisis et al (2007) recorded driving errors, response time, driving speed, and lane position for a full windshield HUD in a driving simulation. Analysis only compared different arrangements of presented HUD imagery and did not directly compare device with small HUDs or HDDs.

Research up to now has only either compared a full windshield HUD to HDDs or examined the full windshield HUD on its own. Findings have not explicitly investigated the boundaries of a driver's FOV, and how a full windshield HUD would affect driver perception and driving performance. Considerations in to human attentional capabilities have yet to be explored when testing or designing AR displays of such size. In the next section I provide an overview of current knowledge of human visual attention.

1.6 Human visual attention

The developments of functional neuroimaging studies and data have supported the existence of three different attentional networks (Fan et al, 2002); alerting, orienting and executive attention (Posner & Rothbart, 2007). Alerting refers to instances of high sensitivity for incoming stimuli; orienting refers to the selection of information from a sensory input; and executive attention

monitors and resolves conflicts which arise amongst thoughts, feelings and responses.

For example in Posner's Attention Model (see *Figure 1.7)* it is hypothesised the three subsets of attention are integrated together. Attention begins with alerting, acting as a 'security guard;' in terms of surveilling the environment. Altering is further broken down in to two states; intrinsic awareness and phase alertness (Posner & Petersen, 1990). The first state (intrinsic awareness) attention is involved in 'monitoring the sensory environment for any unusual activities' (Posner & Petersen, 1990). If something 'unusual' is detected, an 'alarm' is sounded in the brain. Once this happens phase alertness occurs, this is when attention changes from general to specific. This is when attention 'orients' towards the stimulus. Eye movements, or/and head movements are directed towards stimulus. At this stage the executive network is activated. The role of the executive network is to evaluate signals and to make a decision on 'what to do next' (Posner & Petersen, 1990).

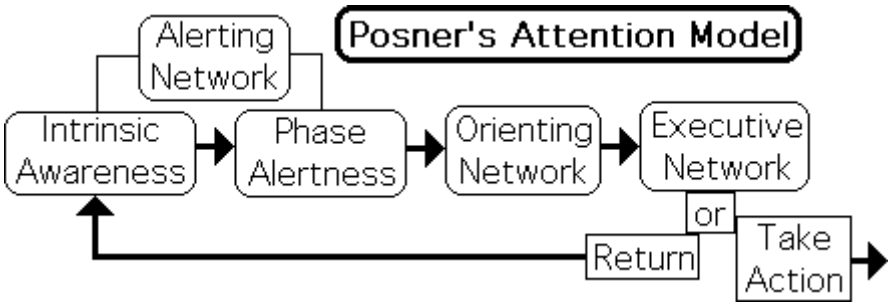


Figure 1.7: Schematic presentation of Posner's Attention Model (Posner 1980)

With the increased size of AR HUDs, more visual information will be available in users' peripheral vision. This may raise issues about distraction and attentional allocation resources; such as task management (Gabbard 2014). Furthermore problems regarding splitting between peripheral and central visual streams brings up unanswered questions and an attentional debate which has

being going for decades: what determines the selective processing of irrelevant information?

Attention defines how humans deal with additional information in a wider field of view. AR devices will be able to present visual information in the periphery of users vision whilst carrying out a difficult foveal task (e.g. keeping in lane when driving). From the visual information of the virtual imagery, and the real world environment users will need to exclude potential distractors and select relevant information.

1.6.1 Early Vs Late selection processing theories

Within attentional theory there has been a long held debate on the processing of irrelevant distractors and the existence of selective processing; an argument of importance as it determines limits of pre-attentional processing. The argument is split between early and late processing mechanisms. An early selection approach (Broadbent, 1958; Treisman, 1969; Treisman & Geffen, 1967) proposed the concept of a limited perception process requiring selective attention to proceed. Early attentional selection is proposed to occur after the analysis of physical features (of a scene/object) and is used to differentiate between selected and nonselected stimuli (see *Figure 1.8*)**;** resulting in unattended stimuli not being fully perceived. In other words humans have a limited capacity for information processing and as a result can only select a small subset of the information for detailed processing (Treisman, 1969). Applying this theory of attention, it may assume that when presenting imagery in peripheral vision it would not be perceived (i.e. in a large full windscreen HUD) because of a perceptual override (too much perceptual information for a user to process), as the central visual task (i.e. of driving) would acquire attentional resources.

The opposing concept proposes that all inputs are analysed preattentively, as attention is allocated depending on current tasks. In this interpretation, attention is required for selection of a response; however attention is not involved for inputs into primary memory (Deutsch & Deutsch, 1963). Late selection theory relies on the concept that a vast amount of semantic processing occurs preattentively; increasing the chance for preconscious processing (see *Figure 1.8).* What this means is that semantic (meaning) analysis is preattentive just as physical analysis is, but it does not mean that every physical stimulus in the visual field is preattentively analysed for meaning (Deutsch & Deutsch, 1963, Norman 1968).

Using the concept of late attentional selection it would assume that users would perceive imagery within a large FOV HUD, as they would be able to process imagery in the periphery and still is able to place attention on the primary task of driving.

There has been considerable amount of evidence supporting both views of selective attention. It has been suggested, however, the difference between the two proposed models has arisen because of differing experimental paradigms implemented; with the shift from early to late selection theories accredited to the increase in complexity of experiments (Kahneman & Treisman 1984).

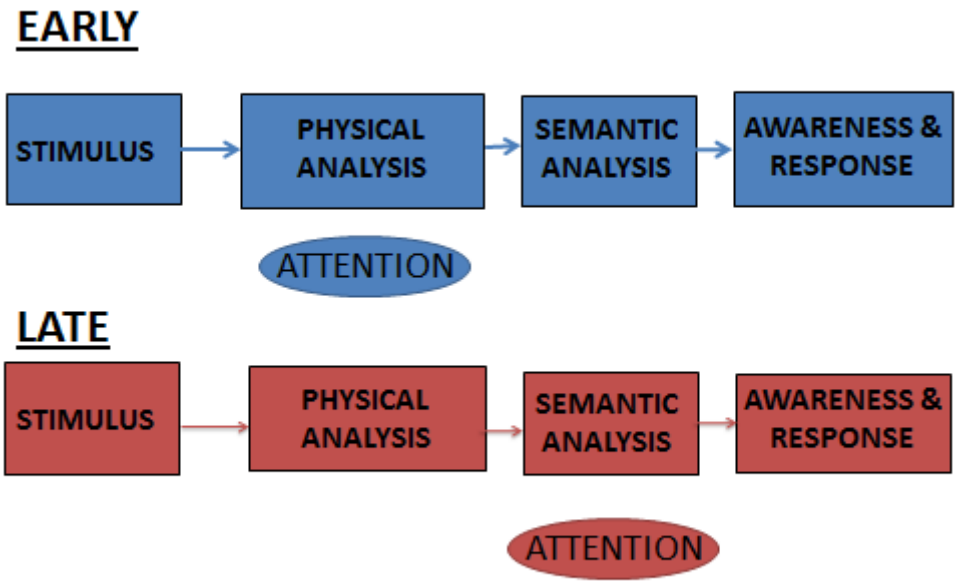


Figure 1.8: A diagram adapted from Pashler (1998) comparing the information processing sequence for both early and late selection theories of attention. The central difference between the two models is highlighted by the point attention is active in the information process; at the physical analysis for early selection and sematic stage for late selection theory.

Early studies supporting early selection were characterised by selective filtering; tasks involving rapid presentation of relevant and irrelevant stimuli usually differing in a particular physical attribute. In a landmark study Sperling (1960) investigated the capability for humans to selectively report items when only viewed for a very brief timescale. Participants were presented with three rows of letters for duration of 50msec (see *Figure 1.9*). A cue was given immediately after the end of display presentation; three different tones corresponding to a different row. Results revealed that when a number of items are briefly presented to a participant, only a limited number are correctly recounted (Sperling, 1960). It was concluded this represents a process limit on information which can be processed.

A number of manipulations, such as delaying the presentation tone, changing cue colours, size and shapes have been carried out over the years (Sperling, 1960; Von Wright 1968, 1970; Turvey & Kravetz, 1970). Findings have demonstrated items are selected more efficiently with a shorter delay of presentation time, the use of colour and brightness, increased size, and that location is improves report compared to shape (Sperling, 1960; Von Wright 1968; Clark, 1969; Turvey & Kravetz, 1970).

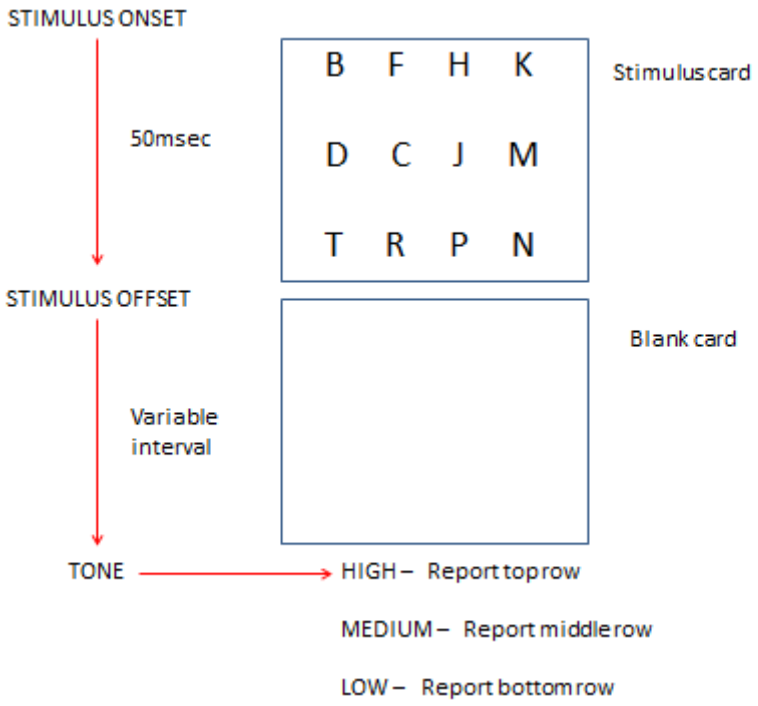


Figure 1.9: Diagram of the sort of display used by Sperling (1960). In whole report the subject reports non-selectively from the display, in partial report the row indicated by the cue is to be reported. In both conditions report is after the display has terminated.

Studies supporting late selection have used more indirect measures of distractor perception in 'selective set' paradigms. In a selective set paradigm stimuli are presented at a slower rate with the participant 'choosing' which of several possible stimuli to expect or search for (Kahneman & Treisman, 1984).

For example, in a Flanker task (a form of selective set initialised by Eriksen and Eriksen, 1974), participants are presented letters, and then required to make a directional response depending on the direction allocated to a specific letter. The participant's task is to respond to a central letter in in a row of five (e.g. SSHSS) as quickly as possible. For instance a right response was allocated to the letters 'H' and 'K,' and left response to 'S' and 'C.' In a flanker experiments there are three types of stimulus which are presented to participants; 'congruent' where the flanker indicates the same directional response as the target, 'incongruent' in which the opposite directional response is indicated from the flanker (see *Figure 1.10*).

The final type of stimulus is referred to as a 'neutral' as the flanker demands neither the same response nor conflict in response. Reaction times were found to be longer for incongruent compared to congruent conditions; referred to as the flanker effect (Eriksen & Eriksen 1974). The results found for incongruent stimuli, automatic activation lead to incorrect, but fast, responses. This required attentional responses to overrule the automatic fast incorrect responses, and resulted in slow correct responses. This was attributed to late-selection as it was suggested as evidence for the activation of attention for semantic analysis. It was concluded attention filters only ensue after meaning is analysed (Eriksen & Eriksen, 1974). Information which is presented is processed to a high degree, where meaning is processed. After meaning is processed the most important items are selected for further processing.

Subsequent variations have implemented numbers, colours and arrows (Lindgren, Stenberg & Rosen, 1996; Rafal, Gershberg, Egly, Ivry, Kingstone & Ro, 1996). Such tasks examined participants' resistance to distraction from relevant (congruent) and irrelevant (incongruent) stimuli, demonstrating that the ability to reject information which is surplus to a participant as a consequence of skilful perception (Neisser & Becklen, 1975; Rock & Gutman, 1981).

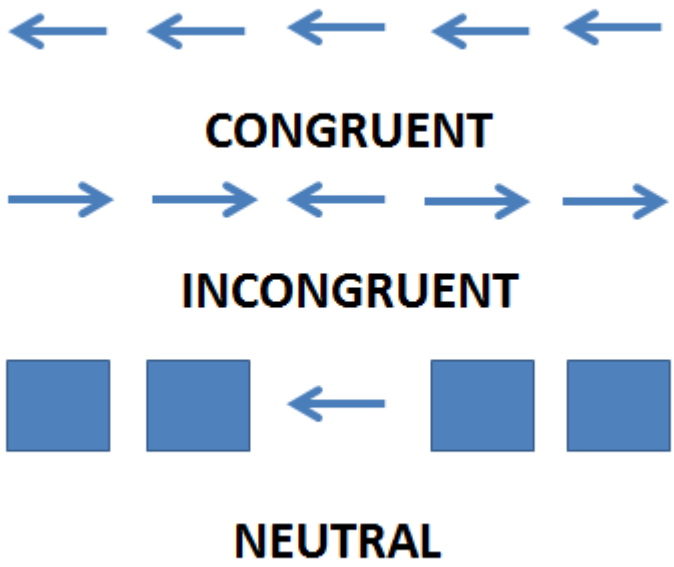


Figure 1.10: A diagram presenting the different types of cues found in selective set paradigms

Thus, studies which support early selection tend to use filtering type tasks, whereas those that support late selection tend to use 'selective set' tasks. A key aspect of studies which support early selection is that participants have to respond from a wide number of choices; often from a complex environment (Kahneman & Treisman, 1984). It has been suggested that the mechanisms supporting tested in the filtering type tasks paradigm are different to those in a selective set experiment; with selective set paradigm resembling automatic processes more than filtering paradigm (Kahneman & Treisman, 1984).

The early-late selection debate may have simplified many aspects of attention and not accounted for a variety of factors involved in the processing of visual displays. Filtering paradigms (as the one carried out by Sperling 1960) involve visual displays where participants have to select one stimulus from a large target display (with a number of stimuli to choose from). Participants carrying out filtering experiments are measured on accuracy where memory

load is high. Results from studies show an early selection of attention, with unattended stimuli going under very limited processing. However the opposite is the case for selective set paradigms (Eriksen and Ericksen, 1974) in which participants select a restricted response from a visual display consisting of only a small set of targets. In these paradigms reaction times are measured where memory load is low. Results from such studies suggest a late selection as attention responds speedily to expected targets.

1.6.2 Perceptual Load and its impact on information processing

Lavie and Tsal (1994) and Lavie (1995) proposed the perceptual load model, which states that under higher perceptual load there is a reduction in distractor interference; indicating perception is more selective under higher perceptual load. This was proposed to address the inconsistencies between the early and late selection models described above.

A number of studies manipulating perceptual load on distractor processing have demonstrated high perceptual load reduces the processing of distractors (Lavie & Fox, 2000; Rees, Frith & Lavie, 1997). Perceptual load is defined as the complexity of the physical stimuli, principally the complexity of the distractor stimuli (Lavie & Tsal, 1994). For instance low perceptual load could be demonstrated by a circle surrounded by triangles. While a circle surrounded by triangles, squares and rectangles would demonstrate high perceptual load.

Lavie (1995) manipulated perceptual load to understand the effect it has on interference for irrelevant elements. Using a variation of the Eriksen flanker task**:** perceptual load was manipulated by set size (from one possible target to six) or by different processing requirements for identical displays; simple detection of target compared to identification of size and location. Results demonstrated that in low load conditions interference effects from a distractor

were significantly more likely to occur than high load. In another experiment a "Go, No-go" situation was carried out. In this situation attentional demands were tested for processing coloured shapes, which were presented next to a target. Participants' response (detection accuracy) (or no response) was dependent on the combination of shape and colour of target. In low load conditions "Go" (respond) was signalled by a blue shape, and not to respond when a red shape was presented. In high load conditions "Go" was only allocated to a red circle or a blue square, and "No-go" (do not respond) by a red square or blue circle. It was predicted by Lavie (1995) that in high load conditions interference from incongruent distractors would be reduced. Results confirmed this assumption as interference was only found in low load conditions.

Results from Lavie (1995) provided evidence of early selection; that attention can prevent distractor perception when the relevant task exhausts perceptual capacity. These experiments also demonstrated late selection occurs in situations of low perceptual load, where irrelevant distractors were perceived in low perceptual load (Rees et al, 1997). Lavie et al (2004) claimed these findings demonstrated the critical impact perceptual load has whether irrelevant distractors are perceived (late selection) or ignored (early section).

In a series of five experiments Lavie, Hirst, de Fockart and Viding (2004) tested out their proposed functions of early and late selection mechanisms. The researchers suggested that the flanker task imposed dual task load on participants as processing distractors in the task depended on both perceptual load and working memory load; and were interested on the effects of switching between attention and memory tasks. To test working memory load on the impact of selective attention, a flanker task (described earlier for Eriksen and Eriksen 1974 design) was presented to participants during a retention interval of short term memory recognition task. On each trial a memory set task was

followed by the selective attention task and a memory probe. Participants' task was to indicate whether the memory probe digit was present or absent in the memory set. In these experiments working memory was manipulated by set size; with low working memory load (1 digit) to high working memory load (6 digits). Conditions of low and high working memory load were compared in terms of distractor capability effects in the selective attention task. In these experiments both working memory (memory set size) and perceptual load (number of letters among which the target had to be searched for in selective attention task*)* were manipulated. In two further experiments in the study, distractor effects in selective attention were compared between single and dual task where participants performed a working memory task.

Results from Lavie et al (2004) found that for task relevant stimuli high perceptual load reduced distractor interference, whereas high load on cognitive control (working memory and task coordination) increased interference from distractors. A passive perceptual selection mechanism enables exclusion of distractors from the early perceptual processes in situations of high perceptual load exhausting the available perceptual capacity for processing relevant stimuli Lavie et al (2004). Also a more active mechanism for cognitive control regulates behaviour in coordination with the current task priorities to minimise invasion from irrelevant stimuli; even in low perceptual load situation in which irrelevant stimuli are clearly perceived. This hybrid model highlights the importance of load in both terms of level and type in selective attention; providing a clear role for cognitive control and perceptual load.

In an AR application, such as a vehicle HUD, perceptual information is used to aid a task which requires large amounts of cognitive control. For instance large amounts of development in the automobile industry are being carried out to produce systems which present navigational information on windscreens to a driver to help reduce the cognitive demand of driving Gabbard

2014). However applying the conclusions presented Lavie et al's (2004) conclusions, in conditions of high driving (cognitive) workload there should be an increased interference of distractors. This argument supports previous assumptions that HUDs only work in low workload (Okabayashi et al, 1992; Gish & Staplin, 1995); is this same for HUDs which demand more of users FOV.

The first experimental section of this thesis (Chapter 2) reports investigations which test Lavie's et al (2004) conclusions on load theory to examine dual task processing involving high cognitive demand (which will be tested with a specific simulated driving task) and perceptual processing on a large display

1.6.3 Cognitive control and perceptual load of attention

Perceptual load theory specifies that human perceptual processing only becomes selective if the perceptual capacity limits are achieved (Lavie 1995, 2005, and 2010). For instance, if a task causes demands to surpass human capacity then items which are deemed as task-irrelevant are not processed; resulting in items being ignored (early selection). However when a task enforces only a small amount of perceptual demand, the capacity that remains is automatically allocated to task-irrelevant items which are processed and can cause distraction (late selection).

Lavie's perceptual load theory of attention and cognitive control (Lavie, 1995, 2005, and 2010) proposes a solution for the debate of early/later selection; both types can ensue depending on the perceptual demands of the current task. According to load theory the capacity of attention is impacted by higher-level 'executive' cognitive mechanisms (e.g. working memory) which determine if distraction occurs; if a task increases in difficulty there is an increase in distractor interference (Lavie, Beck & Konstantinou 2014).

The increased eccentricity of a distractor may also attribute to the complexity of processing stimuli. As discussed above (see 1.3) peripheral vision is not as good as central vision, with poor acuity, making processing more difficult. How has perceptual load theory accounted for information presented in the periphery of the user?

Experiments investigating distractor interference in participants' periphery has found, in conditions of high perceptual load, there is a reduced impact of distractor interference (Lavie, 1995; Lavie & Fox, 2000; Beck & Lavie 2005). For example Lavie (1995) implemented modified response competition tasks, where participants were required to respond to one of two target letters presented in a display. An irrelevant distractor letter was presented in the periphery, to manipulate perceptual load. Specifically perceptual load was manipulated in one experiment by the target letter presented in one of six central positions alone (low load) or amongst five other letters (high load). In another experiment load was manipulated by complexity of target selection; feature detection for low load (e.g. blue or red), or conjunction discrimination for high load (e.g. colour, shape and position). Under high perceptual loads, distractor interference was reduced compared to conditions of low load (Lavie, 1995). However findings from this experiment may not be a result of reduced perceptual processing; rather a late selection effect where the distractors are in fact perceived but the response towards them is inhibited (Lavie, 1995). This is because there would have been a demonstration of greater selectivity (negative priming) under high load if a result of reduced perceptual processing (Lavie, 1995).

Beck & Lavie (2005) implemented a response competition task to compare between distractor processing in the periphery of a display compared to a central fixation point. A response competition task involves participants making a fast response to a target presented amongst distractors e.g. Eriksen

flanker task. Participants were informed to try ignoring the distractors; distractors being classified as either compatible or incompatible with a target response. Incompatible distractors interfere and slow performance; therefore a difference in RTs is compared and assessed with trials for incompatible distractor indicating perception of the identity of distractor.

Over a course of six experiments, Beck & Lavie (2005) findings demonstrated the impact of distractor interference at central fixation location compared to peripheral target. Findings demonstrated that as long as perceptual load of the relevant task is low irrelevant distractors from the periphery are increased. In the high perceptual load task interference from peripheral distractors was reduced, as long as the perceptual load task exhausted attentional capacity (Beck & Lavie, 2005). The authors concluded that the superior processing of foveal vision causes it to be harder to ignore distractors compared to those presented in the periphery. Applying these findings for the use of wide AR displays; if the limits of perceptual capacity are reached then drivers are less likely to detect the imagery presented to them in the periphery.

The interaction with perceptual load was also investigated by Cartwright-Finch & Lavie (2007). The researchers tested a prediction of load theory; that under high load awareness of the task-irrelevant items are likely to be reduced (Lavie, 2005). Within an inattentional blindness paradigm (developed by Mack & Rock, 1998) participants carried out a visual discrimination task; determine which line of a cross was blue (horizontal or vertical line – low load) or which line was longer (high load). The presentation of stimuli was brief (110msec) to inhibit the possibility of eye-movements (results could not be associated to retina blurring of imagery presented). At the end of the final trial participants were asked if they perceived extra stimuli presented in the periphery of the display. Results showed that in the high load condition participants were less likely to report the presence of the peripheral target compared to the low load

condition. However the measure of peripheral detection was retrospective, with the 'unexpected' event occurring after trials and without focussed attention. The experiment may have tested memory related process rather than perceptual ones (e.g. Wolfe, 1999).

The issue of retrospective measuring was rectified by subsequent research from Macdonald & Lavie (2008) who implemented an inattentional blindness paradigm to test peripheral task-irrelevant stimuli. During a central letter task, participants were required to make a response (via button) when they noticed a peripheral stimulus; measuring sensitivity detection. It was found that detection sensitivity was reduced under conditions of high load compared to low load; as predicted by perceptual load theory.

It may be difficult to apply some of these findings to work on AR displays as the sizes of testing displays were only small in size. For instance irrelevant peripheral stimuli were presented 5.4^{o} away from central fixation point in Macdonald & Lavie's (2008) setup, 3.9^{o} for Cartwright-Finch & Lavie (2007) and 3.5^{o} and 4^{o} in Beck & Lavie's (2005) study.

A fundamental issue in visual attention is the relationship between UFOV and eccentricity (Ringer et al, 2016); one of the aims of the thesis is to understand this relationship in more detail by applying attentional theory of cognitive control and perceptual load (Lavie, 1995, 2005). The thesis will explore the impact of perceptual load in a peripheral task on a foveal task.

The impact of cognitive workload (which will be presented in fovea – foveal load) will also be assessed in a split task paradigm. In Chapter 2 a study manipulates cognitive load by using a visually simple virtual simulation which represents a specific driving task; in-lane keeping. Cognitive load was increased by manipulating the size of the 'lane' and difficulty in controlling 'vehicle.' Perceptual load was manipulated by the distance of peripheral target from

participant and the visibility of targets. This approach was developed as AR is currently being deployed in windscreens to try and reduce the already high cognitive demands on drivers by presenting peripheral information to a driver (Rhoda & Thomas, 2006); testing was carried out on a simulated windscreen.

The experiments presented in Chapter 2 will examine if a potential large AR devices (specifically full windscreen AR HUDs) peripheral imagery is ignored (early selection) or distracts (late selection) a user. How do increases in perceptual (peripheral presented) and cognitive (foveal presented) load influence performance on either task; HUD imagery judgment and in-lane driving capability.

1.7 Alternative approaches to peripheral visual field processing

Investigations into the functional area of peripheral vision, during a concurrent foveal task, have found a shrink in the size of peripheral vision; often referred to as 'tunnel vision' or 'attentional tunnelling' (Mackworth, 1965; Wickens & Alexander, 2009). Specifically tunnel vision is referred to as the loss of peripheral vision with preservation of central vision; peripheral vision decreases as a function of eccentricity (Mackworth 1965).

Mackworth (1965) proposed that 'tunnel vision' is a mechanism to prevent the visual system overloading by narrowing the visual field. Mackworth (1965) carried out testing requiring participants to detect similarities between central and peripheral letters. Participants' task was to detect similarities between one letter presented in central vision and two letters presented peripherally (upto 10^{o}). In low difficulty conditions only these three letters were presented (100msec). In high difficulty additional distractor letters were presented in the display. An interaction was found between foveal load and peripheral target

eccentricity; increased levels of foveal difficulty caused a deterioration of peripheral task performance as a function of eccentricity.

Leibowitz and Appelle (1969) examined the impact of a central task on peripheral vision, demonstrating that peripheral discrimination performance is diminished when a foveal (central) task is presented at the same time; with 'tunnel vision' used as an explanation. It was concluded that when central & peripheral stimuli are competing for attention central task is given priority over peripheral tasks (Leibowitz & Appelle, 1969).

Ikeda & Takeuchi (1975) investigated various levels of foveal load and the impact it has on peripheral vision. Their experiment required participants to detect a peripheral target (a star) within an array (upto 10^{o}) of distractors (see *Figure 1.11*) whilst simultaneously carrying out a central identification task (letters or numbers). In the lowest load condition participants would report the presence of two letter (e.g. 'oh,''pq') in the middle condition more complex identification of three letters/digits or a symbol ('sin,''523,'). In the highest load more complex symbols and equations were to be reported ('$\frac{b}{227}$''$4/_2$').

Comparing detection recognition of peripheral stars across foveal loads it was determined that observers could detect peripheral stars at larger eccentricities at low foveal load compared to high foveal load conditions. It was reported that as foveal load increased UFOV became narrower; proposing the effect of tunnel vision. Conclusions from findings of tunnel vision have raised questions on how to interpret results of tunnelling. Is tunnelling the shrinking of UFOV or is it evidence for the attention allocation to the priority of the central task.

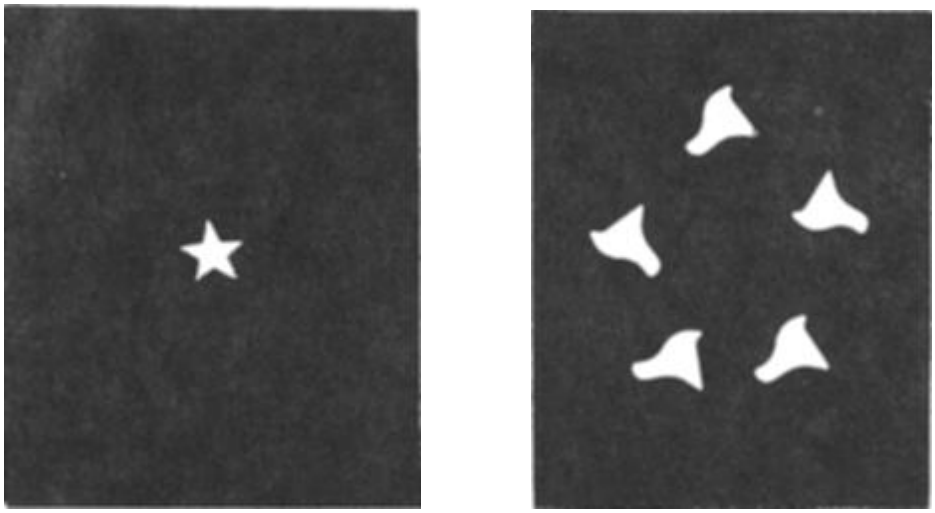


Figure 1.11: Examples of peripheral target (left) and background distractors (right) in Ikeda & Takeuchi (1975) investigation (source: from Ikeda & Takeuchi, 1975, Figure 1 (p. 256)).

1.7.1 Cognitive tunnelling vs general interference

Instead the findings of tunnel vision may (attention is narrower due to foveal load) may actually represent general interference. This refers to when foveal load increases peripheral performance equally declines across all locations in the visual field. In other words foveal load predicts little dependence on visual angle (Holmes, Cohen, Haith & Morrison, 1977).

Holmes et al (1977) investigating the role of task demands affecting UFOV presented both peripheral and foveal geometric forms to participants. The task required participants to report both peripheral and foveal forms or to ignore the foveal form; depending on the condition. Holmes et al (1977) found that the ability to identify a peripheral form was impaired by the presence of a foveal form. However this was not dependent upon the requirement of the foveal form being processed (Holmes et al, 1977).

Gasper et al (2016) examined attentional selectivity by manipulating cognitive load and the impact on UFOV. The experiment involved participants

completing four drives in a car simulator concurrently performing a gaze-contingent Gabor orientation discrimination task. The Gabor patches appeared randomly every 6 to 10 fixations for duration of 0.67s at one of three eccentricities (5 ^{o}, 10 ^{o}, and 15^{o}). Cognitive workload was manipulated in the driving task difficulty (via presence or absence of lateral wind in simulation) and a concurrent auditory N-back task (Kirchner, 1958); that involved determining if letters were repeated. Findings demonstrated Gabor discrimination was significantly impaired across all eccentricities when participants performed the N-back task. Gasper was able to conclude the support of general interference after results showed no difference in the cost of cognitive workload across all peripheral eccentricities (Gasper et al, 2016).

1.7.2 Tunnelling compared to load theory

Conclusions made from research findings on UFOV have been split between tunnel vision and general interference. Initial research and findings were carried out in the 1960s and 1970s; before theories of perceptual load were compiled (Lavie, 1995). Although linking the effects of tunnel vision and general interference with early and late attentional selection have not been explored. The thesis will examine the effects of tunnel vision and general interference by applying attentional theory of perceptual load.

Applying perceptual load theory (Lavie, 1995; Lavie and Tsal, 1994) to findings for UFOV, tunnel vision (Mackworth, 1965; Leibowitz & Appelle, 1969; Ikeda & Takeuchi (1975) can be explained in terms of early selection.

For example in Ikeda & Takeuchi (1975), during low foveal load conditions, where participants had to report the presence of two letters/digits, demand for the task did not surpass human capacity and participants had enough attentional resources to carry out to detect peripheral star, even for when perceptual load was high for detection (i.e. at the furthest eccentricities).

When the foveal task requires more resources for reporting symbols and equations (whilst still attempting to processing peripheral targets) the task begins to demand attention which surpasses capacity of the participant. This results in the secondary task (the peripheral task) items which are deemed irrelevant being ignored. The higher the perceptual load, the peripheral targets furthest away from foveal task, will be ignored.

On the other hand, general interference can be explained in terms of late selection; if the foveal task does not demand enough attention participants could potentially be distracted by peripheral targets. For example in Gasper et al's (2016) experiment when loads for the driving task was low, and did not surpass the demands of the participant, participants could accurately discriminate the orientation of the Gabor patches at all eccentricities in the periphery. However despite when the foveal task increased, the perceptual demand for the peripheral task still remained low, the capacity which remained is allocated to task-irrelevant items and still be able to process imagery in the periphery and focus attention on the foveal task.

Whether participant UFOV experiences early (tunnel vision) or late (general interference) selection has been difficult to conclude for a variety of reasons. Experiments have used a number of different tasks and combinations of tasks (foveal and peripheral); detecting similarities between letters (Mackworth, 1965), discriminating targets amongst distractors (Leibowitz and Appelle, 1969), controlling the illumination of a lightbulb (Ikeda & Takeuchi, 1975).

Williams (1988) proposed three requirements for a demonstration of tunnel vision in a visual task; there must be a sufficiently difficult foveal task, one which is given priority over all other tasks and the participant needs to react fast to both foveal and peripheral tasks. Driving is very similar in this manner; the foveal task of driving and controlling the vehicle is a difficult task, and one

which is given priority over any other task, requiring drivers to react fast to information presented centrally (e.g. vehicles directly in front) and peripherally (e.g. hazards such as pedestrians).

In a set of experiments Williams (1988) assessed the accuracy of identifying single letters presented in the near periphery (upto 4.4^{o}). In the foveal task a pair of letters were presented, participants had to either match the letters in the foveal and peripheral presentations (low load) or match the vowel or constant (high load). Results demonstrated a greater drop off in peripheral letter identification (assessed by correct responses) for hard foveal cognitive load compared to low foveal cognitive load.

Importantly when comparing conditions of dual task and single task (no foveal task), tunnel vision was only found in the dual task condition. This was proposed to as evidence that tunnel vision is based purely on attention; appeared participants had less attentional resources available for peripheral locations (Williams, 1988). According to Williams (1988) the study findings suggested a top-down interference mechanism which is sensitive to both the complexity of the task and the motivation of attentional allocation.

1.7.3 What is the effect of display area?

In the preceding sections I discussed two theoretical frameworks. First I discussed perceptual load theory; attention capacity is limited and is more selective under higher perceptual load with task-relevant stimuli being processed before task-irrelevant (Lavie 1995). Second, I discussed tunnel vision and general interference. In instances of tunnel vision an increase in foveal task difficulty result in peripheral performance deteriorating as a function of eccentricity. In circumstances of general interference when foveal difficulty increases peripheral performance decreases as whole; not specific to eccentricity. Common to both theories is the ideal of high load presented in

central vision and distractors in the periphery. As discussed in section 1.6, few studies have explored a wide range of distracter positions. Rather experiments have presented both peripheral and central task within the same small display. For instance Lavie (1995) presented periphery stimuli upto 2.9^{o} from the central presentation. Williams (1988) displayed periphery targets at 4.4^{o}.

Perceptual load of a stimulus refers to the complexity or difficulty in processing physical stimuli. Few studies have examined distractors far in the periphery (display size above 30^{o}), although the eccentricity of a peripheral distractor could be suggested as requiring more attentional demand. As discussed in *section 1.3* peripheral vision is significantly worse than foveal vision, therefore will require more processing demands.

1.7.3.1 Driving simulator research in UFOV

One area of research which has examined UFOV on large display areas is driving simulator research. Applying past conclusions on foveal load and UFOV Miura (1990) explored the function of UFOV in complex driving scenarios. Participants were tested for responses to peripheral targets on a vehicle windscreen for a variety of roads; with load manipulated by the amount of traffic in a simulated driving experiment. Participants task involved driving safely and verbally responding to the presence of a target stimulus. Results found slower RTs to peripheral targets as the complexity of the driving scene increased.

To account for the result Miura concluded increased scene complexity caused tunnel vision (Miura, 1990). It was claimed as the visual scene complexity increases, due to an increase in the amount of visual information, a person begins to process a smaller proportion of space; resulting in a ‘driver’ allocating attention to information directly in front of them. Building upon tunnel vision theory, in driving scenarios as a scene gets more complex the area

in space a driver directs at becomes smaller, with greater amounts of attention required to process the greater amount of information. It has been hypothesised that a decrease in UFOV is a coping strategy caused by shifting attention from peripheral to central vision (Miura, 1990; Ball, Owsley, Sloane, Roenker, & Bruni, 1992). Supporting evidence by Miura (1990) have been criticised for being dependent upon post hoc analysis, speculative in nature and the difficulty in measuring the 'amount of attention needed.'

Subsequent research from Crundall, Underwood & Chapman (1999, 2002) tested RTs and hit rates of peripheral targets (30^{o}) whilst participants watched video footage of driving scenarios. It was reported that as processing demands increased (manipulated by an increased occurrence of hazards), hit rates for the appearance of visual targets decreased. This finding led the authors to conclude UFOV in driving gets narrower as processing demands increase, a finding of tunnel vision (Crundall, Underwood and Chapman 1999). Conclusions from this study are limited as the experiment involved the passive watching of driving, rather than actively driving.

More recent research from Seya, Nakayashi and Yagi (2013) investigated spatial distribution of UFOV in driving. Testing in a driving simulator, participants were presented either normal or mirror reversed letter E's on driving images in one of 8 directions (in three eccentricities; 2.5^{o}, 5^{o} and 7.5^{o}) from participants eye gaze (horizontal, vertical and diagonal meridians). Participants were required to drive safely and respond to peripheral targets to indicate whether the stimulus was normal or mirror reversed, using a button on the steering wheel. In the experiment a gaze marker was presented before target at the location where participants were looking, moving in synchrony with participants' eye movements Results of the experiment demonstrated that an increase in target eccentricity correlated with an increase in RTs to targets; a demonstration of tunnel vision.

The use of a gaze marker presented at participants' current gaze location before peripheral target stimulus may have affected the shape of participants UFOV. For instance the gaze marker may have distracted drivers, may have been a cue for target stimulus, limitations of eye tracker may have produced positional error between gaze marker and participants gaze. Finally the gaze marker produced small oscillations caused attentional shifts. It was concluded that findings in Seya et al (2013) could be specific to the conditions in the study.

Gasper et al (2016) investigated attentional selectivity for UFOV in a car simulator. Participants completed four drives in a car simulator simultaneously performing a gaze-contingent Gabor orientation discrimination task. Peripheral targets appeared randomly every six to ten fixations for duration of 0.67s at one of three eccentricities (5^{o}, 10^{o}, and 15^{o}). Cognitive workload was manipulated in the driving task difficulty (via presence or absence of lateral wind in simulation) and a concurrent auditory N-back task (Kirchner, 1958). Findings demonstrated Gabor discrimination was significantly impaired across all eccentricities when participants performed the N-back task. Gasper et al (2016) was able to conclude the support of general interference after results showed no difference in the cost of cognitive workload across all peripheral eccentricities (Gasper, 2016).

However Gasper (2016) did not include an explicit foveal load in the experiment, since a foveal load is typically proposed to be necessary for producing tunnel vision/early selection (Williams, 1988, 1989). For instance, experiments' which implement UFOV measures (Ball et al, 1992), or similar methods; involving simultaneous foveal and peripheral tasks, produce findings matching with tunnel vision (Ball et al, 1988; Ball, Owsley, & Beard, 1990; Sekuler, Bennett & Mamelak, 2000).

Driving studies examining UFOV have often implemented gaze-contingent display testing (e.g. Gasper, 2016) or peripheral detection tasks (e.g. Miura, 1990). There are potential limitations to both methodologies. Concerns have been raised how gaze contingent displays create artificial effects on eye movements, such as saccadic inhibition (Reingold & Stampe, 1999). This is when increased fixation durations are caused by late presentations of markers. It has also been argued that covert attention is best measured when participants respond to preselected targets compared to participants selecting items which are targeted by covert attention (Ringer et al, 2016).

The alternative method, peripheral detection tasks, has its own limitations. Firstly target stimuli are not size scaled, thus not controlling for eccentricity dependent sensitivity loss, and secondly the retinal eccentricity of a target is not controlled for either, therefore resulting in measurement noise (Ringer et al, 2016).

Taking into account previous methodologies in UFOV testing, the experiments presented in this chapter will explore further the interaction between a foveal task which uses similar high level cognitive resources as driving, and how the impact of a peripheral detection task (orientation discrimination task in line with Carrasco, 2011; Correa, Lupianez, Milliken, & Tudela, 2004) interacts with increased levels of a foveal task.

Experimental findings on UFOV have shown that humans' ability to detect peripheral targets decreases with increased eccentricity (Gruber et al, 2014). It is also been suggested that increase in foveal load may have detrimental effects on peripheral target detection and discrimination (Seya et al 2013). This could be an issue for vehicle HUDs which aim to command the entire windscreen. Driving is a very demanding cognitive task requiring a lot of attentional resources; so is it possible to have an AR system which projects

information to a driver in their peripheral without consequences of the central task of driving?

It may be difficult, however, to apply previous findings of tunnel vision to foveal tasks which rely on greater attentional demands; such as driving. Early investigations describing a finding of tunnel vision tested on foveal tasks which involve simply memory or verbal recall. Furthermore a vehicle windshield covers a 60^{o} angle to either left or right visual field (depending on the country) of a driver; therefore a total of 120^{o} in size. The experiments presented in Chapter 2 implement a similar psychophysical paradigm to Leibowitz & Appelle, (1969) and Ikeda & Takeuchi (1975) to measure peripheral discrimination performance for a much larger display area (120^{o}).

1.7.3.2 Large displays without driving related task

One example of a study that did use a wide display is the studies of Van de Weijgert (2003) who used a large display for peripheral presentations upto 68^{o} in left and right visual field**.** Two experiments manipulated cognitive foveal load to examine the effects on near and far peripheral visual field. Performance was compared when participants performed a single task, a foveal task with low cognitive load, and foveal task high cognitive load, a peripheral task and two dual task conditions combining the foveal and peripheral tasks simultaneously.

The foveal task was a memory set task; participants were presented with a memory set of one letter (low load) or two letters (high cognitive load). After the set was memorised participants pressed a foot-button and central fixation point appeared. One second after fixation point onset the foveal stimulus (2^{o}) was presented for 150ms. The foveal stimulus either contained one letter from the memory set (positive trial) or no letters from the memory set (negative trial). Participants had to respond with a ‘yes’ or ‘no’ button press. The peripheral task

required participants to identify the location of a peripheral arrow (1.6^o x 1.9^o) presented at eccentricities of 8^o, 20^o, 32^o, 44^o, 56^o, and 68^o.

Participant reaction time analysis revealed an almost equal effect of peripheral detection across all eccentricities. Results demonstrate a suppression of allocating attention (general interference) rather than tunnelling; with increase in cognitive foveal load resulted in general interference on the peripheral identification task (Van de Weijgert, 2003). The findings demonstrated a non-significant difference between central foveal RTs and peripheral RTs.

The findings of Van de Weijgert (2003) were inconsistent with earlier similar studies where participants carried out a combination of a central task (e.g. target tracking, target classification) and a peripheral signal detection/discrimination task at wide viewing angles of upto 30^o (Ikeda & Takeuchi, 1975) and 90^o (Leibowitz & Appelle, 1969). The foveal tasks in early studies include maintaining an illuminated light bulb by pressing a button (Leibowitz & Appelle, 1969) and recalling a list grouped letters.

Jerome (2007) examined three factors which affect the potential performance of AR cues; 1) attentional resources, 2) endogenous/exogenous cuing, and 3) the location of targets outside the FOV. Taking place in a 360^o surrounding AR system, 64 participants were required to search and detect enemies (primary task), for different cue conditions (visual, auditory, tactile), and in two workload conditions (low; 8 targets in 2 minutes high; 16 targets in 2 minutes). A secondary task was included which required participants to 'extinguish' street lights which would light up at various times in the scenario for upto five seconds. Reaction times and detection performance was analysed, with results demonstrating no significant improvement for visual cueing in the AR system. Jerome (2007) attributed the finding to the size of the display used,

a 360^{o} display compared to previous studies using $>30^{o}$, which were claimed not to be wide enough to present stimuli in the periphery of participants. It was concluded that visual cues in the experiment should not have been classified as exogenous, but rather endogenous; as the mechanism appeared to orient attention was top-down in nature rather than bottom-up (Jerome, 2007).

Detection rates and reaction times did improve when all modalities (visual and auditory) combined in a condition, it was proposed participant performance improved because there was a reduction in the amount of parallel information, thus freeing up attentional resources at the perceptual and cognitive stage of information processing, which results in a more salient target (Jerome, 2007).

Analysis did find an effect of workload, interestingly in the high workload condition detection rates were significantly superior to low workload condition. However for the secondary task analysis, RT performance was superior in the low load condition compared to the high load. Results of workload were attributed to a scenario which was not engaging enough with participants bored and inattentive in low load conditions, causing attentional resources being directed to the secondary task. At increased levels of workload participants become more engaged (Jerome, 2007).

Jerome concluded results from the experiment demonstrate bandwidth limitations in the human attention system; only a limited amount of information can be processed from the immense amount which is presented to an individual. Findings were explained through Broadbent's (1958) information processing model with a filter choosing information to further process, with AR cueing acting as an external filter aiding the internal filter for the important information to process.

1.8 Perceptual issues of depth perception in AR

AR systems integrating digital information within a user's environment need the ability to perceptually place virtual imagery at the same point in depth as the environmental object it is highlighting. AR devices which incorrectly display depth (a mismatch in depth between AR image and object), will be less immersive and intuitive, requiring more time to process AR information (Kruijff, Swan & Feiner, 2010). Further consequences of a misappropriate depth match would include inappropriate accommodation and vergence for real world objects (Poitschke, Ablassmeier and Rigoll, 2008); causing systems to be potentially distracting**.**

After examining early research on AR HUDs, with regards to the safety of AR applications, Tufano (1997) strongly suggested that AR imagery needs to be placed at the same perceptual depth as its real-world counterpart. Tufano (1997) further proposed a potential for distraction to the driver if there is a mismatch in terms of depth between a virtual graphic and real-world object; with too much attention allocated to the HUD and away from the primary task of driving.

From an engineering point of view, it has been proclaimed the ultimate goal of AR research is to attain perfect dynamic registration between the real and virtual worlds (Sands, Lawson and Benyon, 2004). Applying psychological concepts and theories of depth perception and cue combination in the real-world the thesis will examine and report depth perception in AR; to aid designers to be able attain improved registration, in terms of depth, between the real and virtual worlds.

Over the last decade there has been a growth of research examining depth perception in AR environments (Livingston et al 2003; Swan II & Livingston,

2006; Swan II, Jones, Kolstad, Livingston & Smallman, 2007; Jones, Swan II, Singh, Kolstad & Ellis, 2008; Jones, Swan II, Singh, & Ellis 2011; Singh, Swan II, Jones & Ellis, 2010; Jerome & Witmer, 2005). Overall research investigating depth perception in AR environments have expressed a trend of underestimation of depth; when virtual imagery in a scene is perceived to be further away than it actually is (Swan II & Livingston, 2006; Swan II et al 2007). However outcomes have been inconsistent, with research methodology and AR device implementation varying across studies (this will be discussed further in 1.12)

After reviewing research in AR perception Kruijff, Swan & Feiner (2010) stated that the most common perceptual difficulty in AR is the incorrect interpretation of depth by users; so why do humans have such apparent difficulty in estimating distance in AR environments? Is this because humans combine cues to depth perception for AR in a different way to real-world depth perception? Or has past research methodologies not produced reliable results because of certain testing parameters? To attain perfect registration between the real and virtual worlds we need discover the depth resolution in AR (the minimum possible distance between two adjacent depth values, expressed in depth distance units of measure).

In the second section of this chapter, details of how humans derive depth in real world environments will be detailed. This is followed by past research investigating real-world depth experimenting examining depth resolution for humans at intermediate distances (1.5 – 30m). The chapter will finish by a critical analysis of depth perception research in AR environments.

1.9 How do humans judge depth

To understand the depth mechanisms in perception of AR, we first must understand how human depth in the real-world works. Depth perception involves a three-dimensional world derived from a two-dimensional retinal image of a scene.

1.9.1 Stereopsis: depth information derived from two eyes

Stereopsis is the 3-dimensional perception of the world derived from visual information attained from two eyes; binocular vision (animals with two eyes producing a single 3d image). The two eyes are located at different horizontal positions on the head, producing two different projections to the retina. The positional difference between the eyes is called binocular disparity. Depth perception is derived from the visual cortex processing these disparities. Although depth perception can be achieved through the use of one eye; visual information such as relative size, brightness or motion parallax, judgments are superior when obtained from binocular disparity (Palmisano et al, 2009).

To maintain single binocular vision both eyes are required to move simultaneously in opposite directions (vergence); when focussing on an object both eyes rotate on a vertical axis allowing for a projection in the center of both retinas. For an object near an observer both eyes will rotate towards each other (convergence), and rotate away from each other to focus on an object in the distance (divergence). Another mechanism used in focusing on objects in depth is accommodation; ciliary muscles of the eye relax (focussing on distant object) or contract (focussing on near object) to maintain clear focus in depth (Helmholtz, 1855). Not all humans can use stereopsis for depth perception, as approximately 10% are stereoblind (Blake & Hirsch, 1975).

1.9.2 Optimal Cue Integration

To achieve a three-dimensional perception of the world depth cues are used to disambiguate depth relationships into reliable precepts' (Kaufman, 1974). When multiple cues are available to an observer it is important to understand how observers combine these cues together.

1.9.2.1 Reliability and optimal cue combination

To efficiently combine information provided by multiple cues an observer must calculate the reliability against the unreliability of the information provided (Jacobs, 2002). One way of interpreting reliability is in terms of ambiguity, with every cue derived from the visual environment being ambiguous; with reliability being measured in terms of variance (Knill & Pouget, 2004).

A reliable source of sensory information is one which produces a small amount of error from a signal. Applying probability theory to cue integration, it proposes both an inherent uncertainty available to our sensory system (stochastic error), and the uncertainty of encoding that information (erroneous calculation). This results in the neural representations of sensory information guiding behaviour responses in a probabilistic manner (Angelaki, Gu & DeAngelis 2009).

For example optimal cue integration theory (See Figure 1.11) uses Bayesian predictions for relationships between probability distributions and behavioural response functions.

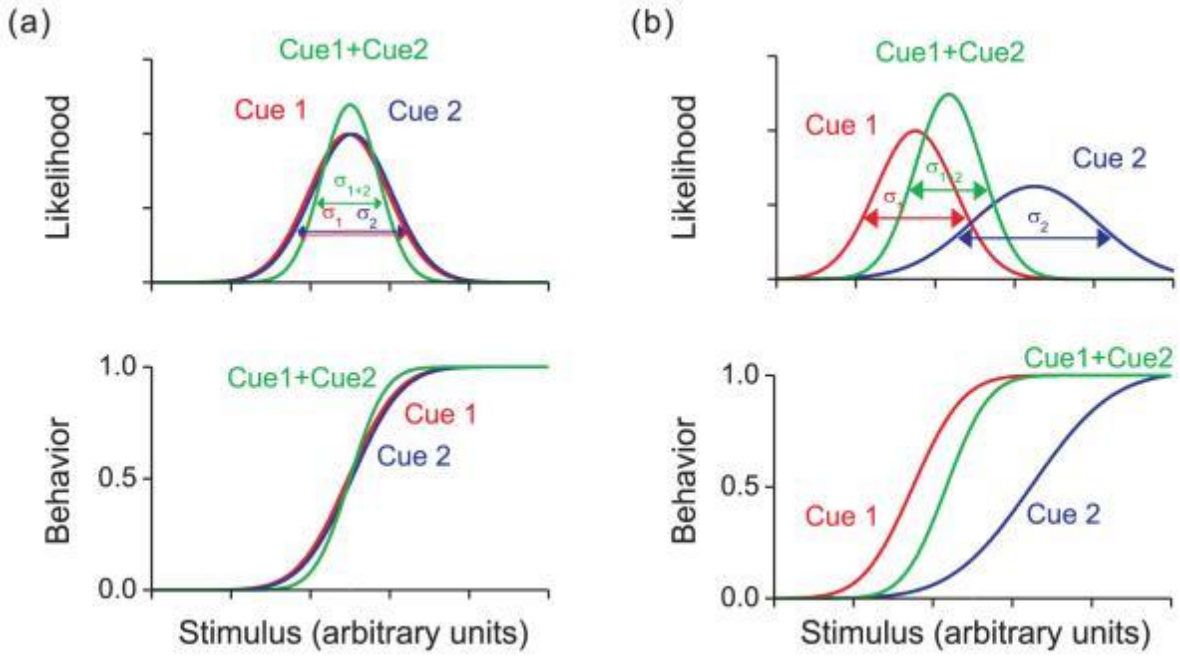


Figure 1.12: Bayesian predictions for the between probability distributions (top) and behavioural response functions (bottom). Green curves represent the situation in which two sensory cues are presented together, whereas the red and blue curves represent the relevant quantities for each cue presented separately. In (a), cues 1 and 2 have equal reliability (same variance), where as in (b) cue 2 has a larger variance (lower reliability) than cue 1. (Source: from *Angelaki, Gu & DeAngelis, 2009, Figure 1).*

For example optimal cue integration theory (See Figure 1.12) uses Bayesian predictions for relationships between probability distributions and behavioural response functions. In Figure 1.12 green curves represent a situation where two sensory cues are presented and combined together. The blue and red curves represent the individual cues presented separately. In Figure (a) the two cues are understood to have equal reliability; in that they have the same variance. Therefore the probability distribution is equally influenced by both cues. In Figure 1.12 (b), cue 2 has a larger variance (lower reliability) than the other cue. The weight of a cue can be determined by its variance. According to

Bayes' rule a posterior distribution (green curve) has smaller variance than the likelihoods for each cue (red, blue) and therefore its mean will be closer to the cue with higher reliability (red) (Angelaki, Gu & DeAngelis 2009).

1.9.2.2 Reliability and viewing geometry

The reliability of cues which provide an observer's visual depth information may vary depending on viewing geometry. The reliability of cues which provide an observer's visual depth information may vary depending on viewing geometry. The minimum difference that an observer can perceive is dependent upon on the amount of noise or variability both found in the signal and the brain. Therefore resolution does depend on variance. For instance across viewing distances the reliability of cue information differs, and fluctuates depending on each particular cue. Cutting (1997) grouped together depth cues into pictorial, kinetic, physiological and binocular disparity (Cutting 1997). Pictorial depth cues are those which can be observed in photos/pictures which give the impression of multiple depths; occlusion, height in visual field, relative size, aerial perspective, relative density, relative brightness and shadows. Kinetic cues provide information by changing the observer's perspective; motion parallax and motion perspective. The muscular control of the eyes provides physiological depth cues; accommodation and convergence. Binocular disparity combines the two horizontally offset views of a scene provided by the eyes. Over different distances depth cues vary over their salience and effectiveness.

Cutting and Vishton (1995) divided perceptual space into three distinct regions; personal, action and vista. Personal space extends to 1.5m (slightly beyond an arm's reach), as depth perception acts almost veridically within this range. Action space is between 1.5m and 30m, for stationary observers, this can

become compressed (items become closer than they really are). Vista space goes beyond 30m to optical infinity. In this region, as distance increases, depth perception becomes increasingly compressed. Within each of these regions, depth cues vary in their availability, salience and effectiveness, determined on perceived reliability (see Figure 1.12).

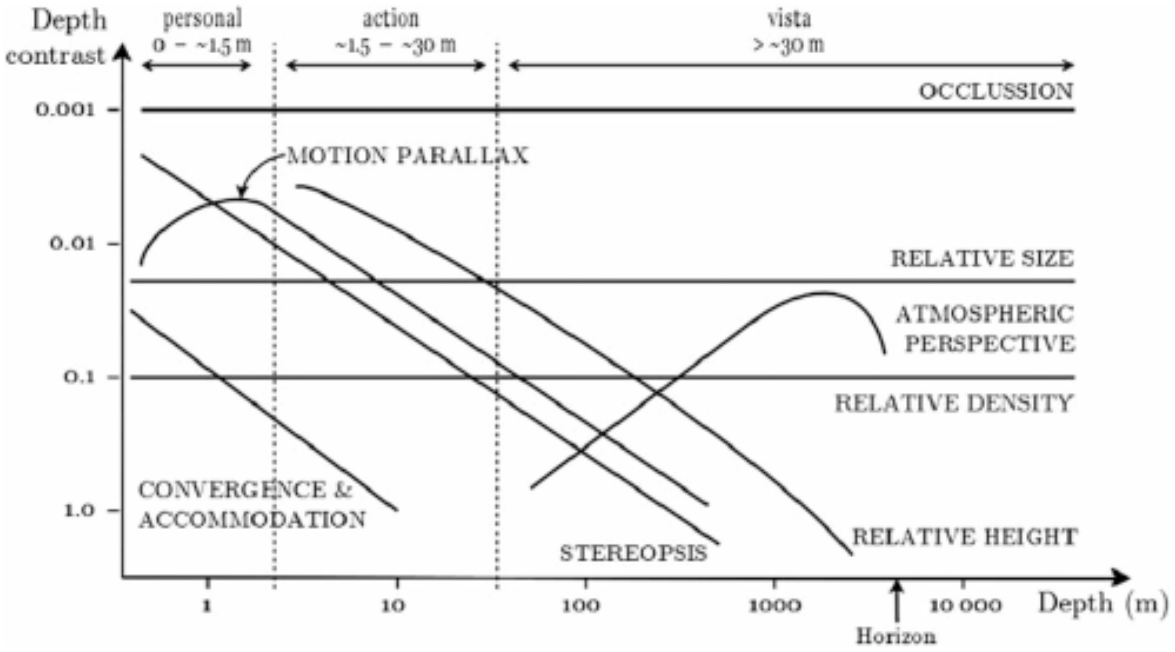


Figure 1.13: Cutting and Vishton (1995) categorized the effectiveness of monoscopic and stereoscopic depth perception cues across three spaces: personal, action and vista. (Source: from Cutting & Vishton, 1995, Figure 1 (p.80)).

The reliability of depth cue information is impacted by how change in the signal is measured, i.e. the error in the measurement. Human depth resolution is defined as the range of depth over which a signal changes (increase or decrease) by a specified amount when profiling between two stimuli (Hofmann, 1976). For example one measure of resolution is the difference in relative depth that corresponds to a change in performance.

From what we know about the human visual system when two objects are far away in the distance, it is likely that they can be relatively far apart and still appear to be at the same distance (Campbell and Green, 1965). When two objects are nearer, the resolution of the visual system is likely to be improved. Specifically two objects are likely to be seen as detached at smaller separations when the object is near compared to when the object is further away; as found in real-world depth judgments research (Palmisano, Gillam, Govan, Allison, Harris, 2010; Allison, Gilliam & Vecellio (2009).

The Weber-Fechner law proposes a concept of human perception where the resolution of perception diminishes for stimuli of greater magnitude from the observer (Fechner, 1860; Fancher, 1996). This is expressed when a JND (just noticeable difference) between two stimuli is proportional to the magnitude of the stimuli. Assessing the quality of a signal does not just involve measuring the reliability of a particular cue or cues; it also involves examining an observer's sensitivity to change in behavioural response as there is a change in stimulus intensity.

The JND is a measure for gaining a differential threshold, where the level at which an increase in a detected stimulus is perceived (Weber). For example in depth perception a low JND threshold indicates a good ability to discriminate between two objects at separate depths with only a small distance between them. Whereas a high JND threshold indicates a poor ability to discriminate between two objects with small spate depths, and would require larger distances between two objects to identify them as separate. This measure of sensitivity for the ability to discriminate is assessed by the slope of the psychometric function; steep slopes represent low threshold (high sensitivity), and shallow slopes represent high thresholds. The slope of a function represents the precision of a response, or in other words the sensitivity to stimulation. For example a steep

slope indicates only a small change in distance and could result in a larger behavioural response.

1.9.2.3 Measuring depth resolution and obtaining PSE

In an experiment measuring depth resolution typically an observer is presented with two stimuli, a test and standard stimulus. The task involves a standard stimulus placed at a specific distance (e.g. 5m). A test stimulus would also be placed in reference to the standard stimulus either infront or behind, with an observers task being to judge if the test stimulus is either in front or behind the standard stimulus.

A two alternative forced choice (2 AFC) paradigm can produce a JND threshold. In this case, the threshold indicates the ability an observer has to discriminate between the position of two objects.. An observer's point of subjective equality (PSE) calculates when two stimuli (test & standard) are perceived to be subjectively the same, and therefore an observer would pick randomly between them**.** The PSE is equal to an x-axis value associated with 50% of the response on the y-axis (0.5 probability point). The PSE gives a measure of bias (an inaccuracy in judgment), in the case for depth perception this allows for an estimate of under or overestimation of depth. An underestimation of depth is when an observer would perceive the test stimulus to be closer to them then they perceive. Whereas an overestimation of depth an observer would judge the test stimulus to be further away from the observer than it actually is (for example see Figure 1.14).

An obtained JND threshold can also be analysed in terms of a goodness-of-fit statistic to assess variability. The statistic ('r' statistic) reveals the extent to which data points (or error bars) fall near a psychometric function. If 'r' is statistically greater than that is expected by chance ($p < .05$), then it is assumed

confidence can be derived from the estimates for thresholds and PSE. Importantly ‘r’ statistic can give a measure of precision, which corresponds to the fineness of a measure e.g. over a fine amounts units of distance there is a big change in a behavioural response. Therefore two obtained functions can be similar in precision but differ in accuracy.

1.9.2.4 Signal in noise & ideal observer

There is variability in the estimates derived from an observer, both in terms of the strength of a neural signal and the noise in the visual nervous system. These both limit an observer’s ability to perceive. Humans have an internal variability, or noise (Barlow, 1957) which is trial-by-trial variability in nature, with the same stimulus presented producing a different response on every presentation. Rather a distribution of responses is produced across identical stimulus presentations. The magnitude of internal noise will determine the variation of the response distribution. Internal noise can be produced from stochastic properties of sensory neurons (Croner, Purpura, & Kaplan, 1993) to fluctuations in participant strategy (Burgess, 1990).

To recognise the potential factors which limit visual performance a human observer is compared to an ideal observer; a machine which knows the stimulus perfectly and exactly, with no internal noise. In the instance of an ideal observer, every time the same stimulus is presented the same decision is made.

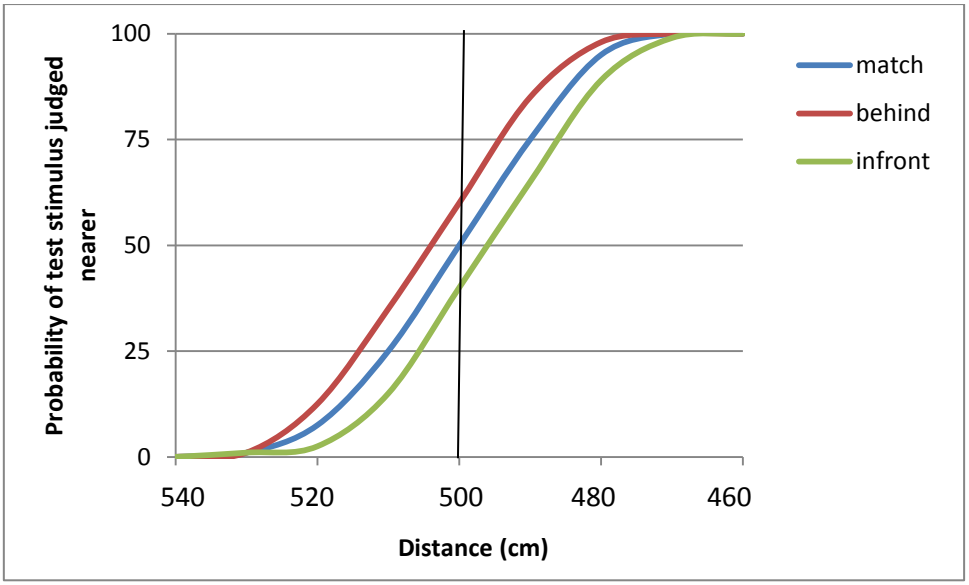


Figure 1.14: A graph representing three example psychometric functions for a depth matching task with a standard stimuli distance of 5m (x-axis). The vertical black line at the 5m point represents the distance of a standard stimulus. The y-axis represents an observer's response probability that the test stimulus is judged closer than the standard stimulus. Psychometric functions are represented by three coloured curves. The PSE is calculated at the 50% value for observer responses to judge the test stimulus to be closer in comparison to the standard stimulus. For example the red curve represents a condition of 'underestimation,' where an observer's PSE requires the test stimulus to be placed behind (in terms of distance) to match the depth position of the standard stimulus. In the example the test stimulus would need to be placed at 505cm to been perceived at 500cm. In this example the 75% point on the curve illustrates the point in depth where an observer will perceive the test stimulus to be in front of the standard stimulus (for the red curve that is 495cm). In the example the green curve represents a condition of 'overestimation,' where the PSE requires the test stimulus to be placed in front of the standard stimulus, for both stimuli to appear at the same distance. In this example the test stimulus would need to be placed at 495cm to be perceived at 500cm. The blue curve indicates

that an observer perceived a test stimulus to be at an equal distance to the standard stimulus.

1.9.2.5 Models of optimal cue combination

A popular model of cue combination is 'winner takes all' where the 'best' cue from an input is selected and everything else ignored (Grossberg, 1973; Yuille & Grzywacz, 1989). The decision from the perceiver is to determine which is the 'best' cue? For example if size is a more effective cue in a far-away scene, it would be classified as the 'best' cue. The combination of cues could be defined on a case-by-case situation; where of two cues (i.e. size and brightness) we select the one with the least variability. If this is the case in real-world depth perception, it may not be the case for an AR scene, where cues found (available) in the real world environment may not be found in an AR environment (Landy & Maloney 1998).

There are further proposed ways different cues could combine to contribute to the perceived distance of an object/input. For instance using unweighted averaging both size and brightness depth cues could have an equal impact to the perceived distance location, despite the 'quality' of the cue; with the location of distance being judged equidistance between the individual cues. On the other hand, if the observer does not perceive the distance to be equidistance between both cues, it suggests the reliability of the cues contributes to the weighing of each cue. In this instance the brain could be performing a form of 'maximum likelihood estimation (MLE) to estimate the depth location. To form these weights in a MLE model, the variability of the localisation judgements of each cue are used, then the contribution of each cue is used to make the depth judgment (Johansen & Juselius, 1994).

Another approach for combining sensory cues applies Bayes theorem; proposing perceptual computations are optimally combined (Edwards, Lindman & Savage, 1963). Humans combine perceptual data about the presence of a signal (likelihood) with a previous expectation of if that signal was going to be present (prior). Previous occurrences of a signal form prior information. Prior information then is combined with other cues in the scene to create a posterior probability distribution. This involves the mean and variance as representing of the most probable stimulus and the probability that the stimulus present, and could represent the explicit percept which is perceived (Sharman, 2014). The weight of a given prior will depend on the ambiguity of sensory inputs. Kersten & Yuille (2003) proposed that higher perceptual uncertainty will result in more weight given to prior information than then sensory inputs. In a Bayesian model, all levels of a computation involves representations of all parameters and probabilities, allowing for cues to be integrated without having to make a commitment early on in an interpretation.

In the opposing MLE system there is no presumed prior information that will affect a decision. MLE argues that weight means of the signal are only determined by their variance and not previous statistical history.

The visual system (Saunders & Knill 2004) must always have an available internal representation of current uncertainty which is flexible in response to novel information. Saunders & Knill (2004), measured slant discrimination thresholds for stereo and texture cue weights. Participants were presented with a sequence of paired surfaces, the task was to judge which surface was more slanted (angled). There were three conditions of surface; texture only, stereo only and stereo and texture. Using a Bayesian theory prediction found the relative weights of different signals are dependent on the expected sensory noise associated with signals. Also the researchers demonstrated the visual systems' ability to adapt to uncertain changes in an

environment (Saunders & Knill, 2004). This was demonstrated by the finding that artificial noise can be used to manipulate the reliance on cues; suggesting a model of uncertainty of available at all times.

The visual systems capability to adapt to different environments has been taken into consideration when modelling depth perception. In environments where there are a small number of uninformative cues to depth, resulting perception may depend on heuristic prior constraints, as well those depth cues which are available (Landy & Maloney, 2000). Reviewing over a decade of work on depth perception Landy & Maloney developed the Modified Weak Fusion (MWF) to account for both cue rich and cue poor environments; based on a linear cue combination, cue weights which are based on reliability and availability, with an interaction of cues for the purpose of cue promotion (depth cues are scaled and then disambiguated before being averaged) and cue weight modification to form a robust cue combination rule. As cue weights change from varying scenes, in environments where cues to depth are less stable, reliable and accurate (such as in VR and AR) priors constraints are used to disambiguate depth stimuli (Landy & Maloney, 2000)

Reviewing previous work to develop their theory, Landy & Maloney only analysed psychophysically tested depth in real-world environments (either cue rich or cue poor). But how do cues combine in virtual and augmented reality? Would the MWF model be able to account for depth perception in such environments? In the present thesis data was collected in an AR environment to gain further understanding to how cues to depth interact and combine.

1.10 Human intermediate distance perception in the real-world

How good are people at perceiving distance in the natural environment? There has been a large spread of studies examining how people perceive both egocentric (absolute distance from the observer) and exocentric distance (the relative distance between two objects) in real-world environments. The general finding is observers underestimate egocentric distance with increasing distance (Gilinsky, 1951; Harway, 1963). In a visually guided judgments task Loomis et al, (1992) asked observers to match a ground positioned depth interval with an interval placed in the frontal plane. It was reported that observers consistently perceived the same interval as shorter in the ground plane than in the frontal plane (Loomis, Da, Silva, Fuijta & Fukusima, 1992). However experiments which implemented an action directed task (such as triangulated walking) to judge depth have found observers are actually very capable at acting on distance (Philbeck & Loomis, 1997; Riesser, Ashmead, Talor & Youngquist, 1990). The way in which depth is measured has large impact on observer's perceived depth (see *Table 1.3* and *Table 1.4*). In visually directed experiments observers are required to blind walk to a previously perceived target; usually up to 20m (action space). After 20m there is a trend for observers to underestimate depth in the real-world (Palmisano, Gillam, Govan, Allison & Harris (2010)

Despite the vast theoretical development and testing on human depth perception, there have only been a few studies which examine depth perception at distances beyond a few meters. Most empirical investigations beyond 2m have concentrated on depth discrimination simulated via vergence in a stereoscope or haploscope (Amigo, 1963; Dees, 1966; Kaufman et al, 2006). Research in clinical and applied areas has typically tested stereoacuity at 3 to 6m (Adams et al, 2005; Kaye et al, 1999; Wong, Woods & Peli, 2002). Most experimental work on stereoscopic depth discrimination has tested perceived

depth magnitudes no further than 4m (Loomis et al, 1992; Drugin, Proffitt, Olson & Reinke, 1995).

Investigations for real depth estimates have been made at 30m, 17m and 6m (Crannell & Peters, 1970; Hirsch & Weymouth, 1948; Howard, 1919), with researchers finding binocular improvements (superior depth judgments with two eyes) when available monocular cues are weak. With salient monocular cues binocular improvements are more uncertain (Teichner, Kobricj & Dusek, 1956 to 30m, Jameson & Hurvich, 1959 to 48m, Teichner, Kobrick, & Wehrkamp, 1955 at 30m to 914m).

1.10.1 Binocular disparity at intermediate distance testing

Using natural scenes Cormack (1984) was the first, and till more recently the only study of perceived depth using disparity at long distances. The observers' task was to set a probe to the depth of another object. The results only revealed that observers can depth match at large distances in the real world, but gives no insight into perceived depth at large distances.

Allison, Gilliam & Vecellio (2009) measured stereoscopic accuracy and precision judgments of the magnitude of depth intervals at longer distance (up to 18m); allowing for a baseline stereoscopic performance for depth estimations. In a series of experiments observers were required to judge the relative depth of a target (black steel rod) with respect to a reference surface (architectural panel). Stimuli were designed so observers would not be able to use the relative size to make judgments. The observers' task was to judge if the test rod was located in front or behind the architectural reference panel. An adaptive staircase method was applied to gain depth discrimination thresholds. Three experiments were set up to 1) compare the performance between binocular and monocular depth perception 2) explore both accuracy and

precision of depth interval estimations and 3) to examine the extent rich natural distance cues improved stereo depth constancy.

Testing was completed at 4.5, 9.0 and 18.0m (distance of the reference surface). Findings from the study demonstrated lower (better) discrimination thresholds for binocular vision compared to monocular viewing (as a factor of 10). Binocular depth estimates were found to be larger and less variable than monocular estimates; with monocular conditions perceived depth was significantly compressed compared to binocular.

The researchers concluded that stereoscopic does work at modest to large distance (upto 20m at least), and is the cause for binocular improvement. Finally they discovered that monocular and binocular distance cues (such as a ground plane) encouraged stereoscopic depth constancy for testing distances of 4.5 and 9m. For real-world environments it was concluded that interactions between depth cues and stereopsis was not limited to improved precision of depth judgments. However, also distance information derived from other sources, such as perspective, can be effective in depth scaling from disparity. In the rich natural cue condition it was depth constancy which was improved. Depth constancy is a process which compensates for the magnitude of retinal disparity on an observation distance, making depth perception veridical (Wallach, Gillam & Cardillo, 1979)

The result further highlighting the impact the background environment has on depth perception; raising the question whether humans take the majority of depth information from the object being judged or the peripheral information derived from the scene? In an augmented reality scene peripheral depth information may have a greater impact on distance judgements due to the variability of depth information derived from a virtual object.

Building on the work of Allison et al (2009), Palmisano, Gillam, Govan, Allison & Harris (2010) explored stereoscopic depth perception at very large observation distances of upto 248m. Adapting on the experimental method of Allison et al (2009), experimentation took place in railway tunnel, for pairs of LED targets presented in a well-lit environment (as far as the nearest LED), as well as an environment in darkness. Results demonstrated improved performance for stereopsis compared monocular viewing at 20 and 40m observation distances, as well as almost all target separations (including 248m). It was concluded that humans use perspective and relative size at far ranges to form a perception of depth.

These experiments of long distance perception both reveal the large impact stereo viewing has at long distances compared to monocular viewing. Both experiments were conducted in environments which already induce linear perspective; railway tunnels and hallways. It is well understood perspective improves depth perception, especially at far distances (Cutting & Vishton, 1995).

In both studies large individual differences in magnitudes of depth estimates were reported, this has been previously found in depth perception studies in AR (Swan II et al, 2007). There would be an expectation of larger individual differences in AR than real-world due to more unstable percept derived from a virtual image (Cutting, 1997). It is vital to know what the extent of variability in terms of depth judgments for AR environments is. If variability is small then it can be assumed systems presenting depth could be applied for the majority of users. However if depth perception varies vastly from one observer to the next, systems may need to be calibrated individually for each user for virtual images to match the real-world object they are highlighting.

Experimental methodologies of previous research in real world depth perception (Allison et al, 2009; Palmisano et al, 2010), were applied for the current experimental paradigm to gain stereoscopic depth estimates in an AR environment. Previous research in AR has rarely tested depth perception at distances upto 25m without the use of a HMD, and applying a perceptual matching task (see 1.11).

The aim of Chapter 3 Section A was to accurately assess stereoscopic depth perception for an AR environment using the same (or as similar) methods to real-world assessments of stereoscopic depth perception. Therefore comparisons between AR and real world depth perception can be more accurate. All the experiments conducted and presented in Chapter 3 implemented a perceptual matching task procedure.

1.11 How does depth perception differ in augmented environments?

There have been a number of studies examining distance perception (both egocentric and exocentric) in virtual environments; with a general suggestion that humans underestimate depth in VR (Willemsen and Gooch, 2002: Loomis & Knapp, 2003; Thompson et al, 2004; Livingston, et al 2005; Swan II et al, 2007).

For example Witmer and Sadowski (1998) examined visually directed walking, in a corridor for real world targets and compared with results to directed walking (on a treadmill) and verbal judgments for a virtual target within a simulated corridor (5 – 35m). Witmer & Kline (1998) carried out the same testing using verbal reports. It was revealed that distances in virtual environments were underestimated compared to those for the real-world environment for both visually directed walking and verbal report (see *Figure 1.15*).

Study	Distances (m)	Method	Accuracy
Witmer and Sadowski (1998)	4.6 – 32.0	Treadmill Walking	92%
Knapp (1999)	5.0 – 15.0	Triangulated Walking	100%
Willemsen et al. (2004)	4.0 – 8.0	Triangulated Walking	88%
Willemsen et al. (2004)	4.0 – 8.0	Direct Walking	100%
Willemsen and Gooch (2002)	2.0 – 5.0	Direct Walking	100%
Thompson et al. (2004)	5.0 – 15	Triangulated Walking	95%
Interrante et al (2006)	3.0 – 9.2	Direct Walking	100%
Swan II et al. (2007)	3.0 – 7.0	Direct Walking	88%
Jones et al (2011)	3.0 – 7.0	Direct Walking	94%

Table 1.3: A table displaying previous studies findings (in terms of distance accuracy), distances tested and method used for real world depth perception (source: from Jones et al 2008, Table 2 (p.14)).

Using a similar methodology as Witmer and Sadowski (1998) and Witmer and Kline (1998) across two experiments Knapp (1999) placed virtual spheres on a textured ground plane at distances of 1, 2, 2.5, 4, 6, and 18m). As well as using verbal report and visually directed walking, a judgment was based on perceived sized. This involved participants turning a knob which changed the width of a vertically oriented aperture; the task was to make the aperture appear just passable (as wide as observers' body).

Study	Distances (m)	Method	Accuracy
Willemsen et al. 2004	4.0 – 8.0	Triangulated Walking	67%
Willemsen et al. 2004	4.0 – 8.0	Direct Walking	85%
Swan II et al (2007)	3.0 – 7.0	Direct Walking	92%
Jones et al (2011)	3.0 – 7.0	Direct Walking	83%

Table 1.4: A table displaying previous studies findings (in terms of distance accuracy), distances tested and method used for real world depth perception viewed through a HMD (source: from Jones et al 2008, Table 2 (p.14)).

The available cues to depth were binocular disparity, convergence, texture gradient, linear perspective and angular evaluation. It was found that across all measure distance was underestimated (see *Figure 1.14*).

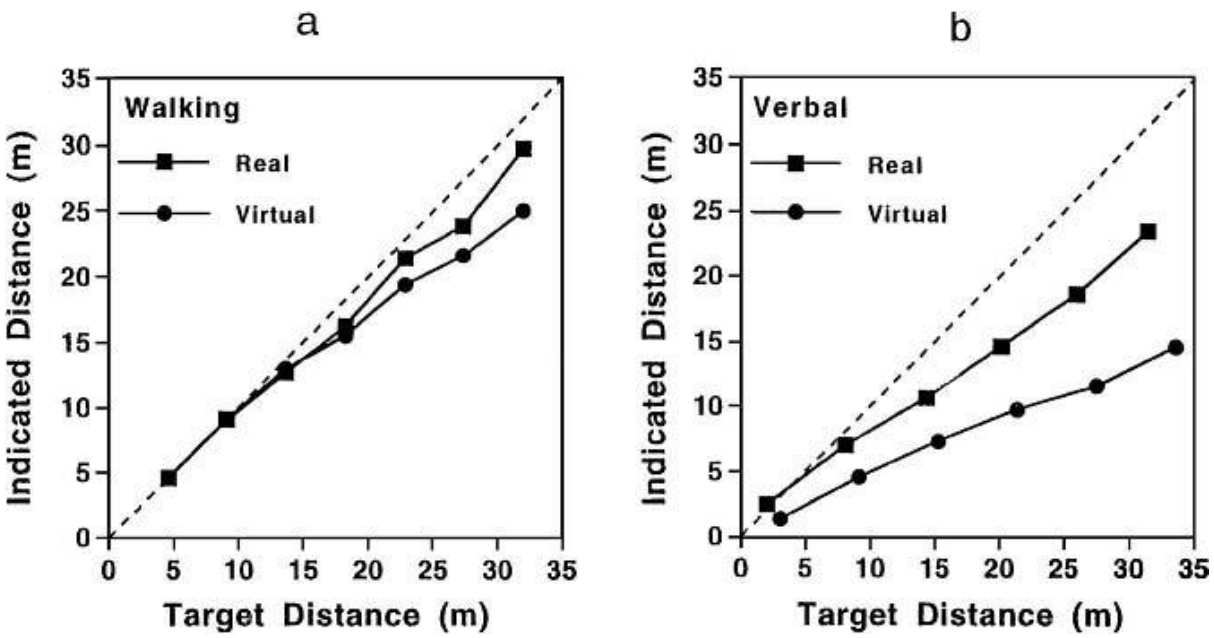


Figure 1.15: Resulting representing (a) distance measured by visually directed walking in real and virtual environments (source: Fig. 2 from Witmer & Sadowski, 1998, (p. 483)). (b) Distance as measured by verbal report (source: Fig. 3 from Witmer & Kline, 1998, (p. 155))

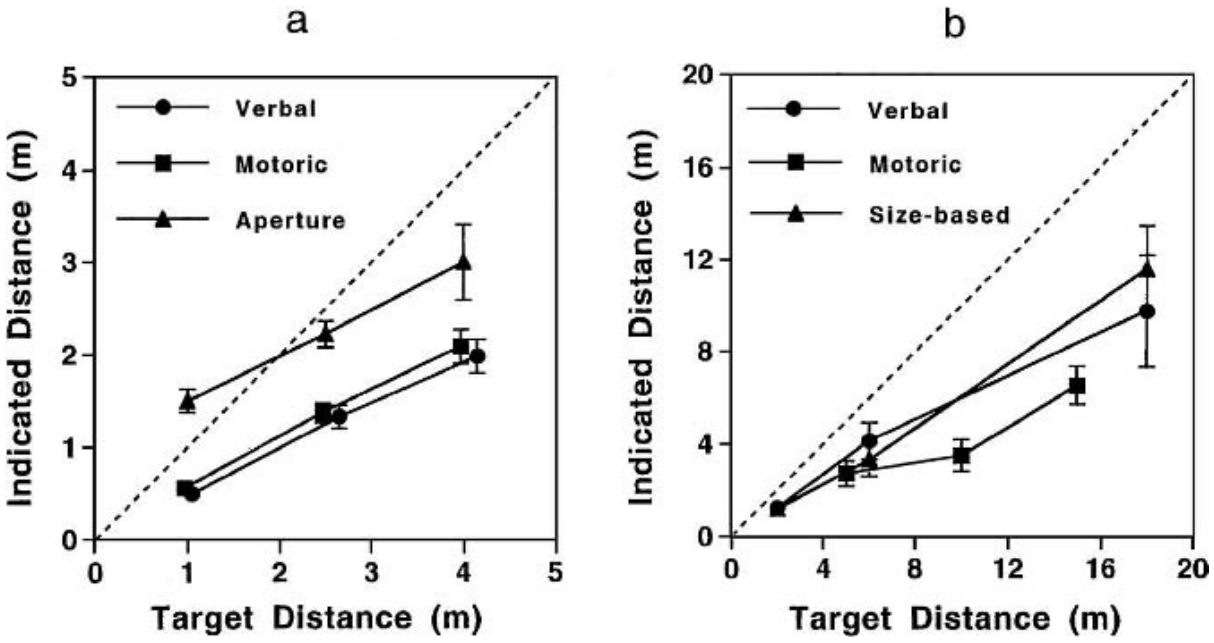


Figure 1.16: Results representing distance judgments comparing (a) three measures of egocentric distance (verbal report, walked distance and aperture adjustment) and (b) three measures of egocentric distance (verbal report, walked distance and perceived size judgment (source: Fig 2.7 from Loomis & Knapp 2003 (p.38)). Error bars represent 1 SEM.

The underestimation of depth in VR has been attributed to the use of limited FOV due to the wearing of HMDs and the poor rendering of graphics in systems. However research has been found to down play the importance of issues for VR depth perception (Loomis & Knapp, 2004; Thompson et al, 2004).

Research indicates a discrepancy between VR and real-world depth perception, however the nature is unclear and leads to predicting AR depth perception difficult. The same effect of underestimation and error has been found in AR research. Over years of research AR underestimation has been accredited to limited FOV of devices, monocular vs stereo viewing, the weight and inertia of the HMD, the effects of implicit feedback and the quality of the

rendered graphics. The following sections of the thesis explain the possibly reasons for underestimation which has been found in virtual environments.

1.11.1 How depth has been measured in VR and AR

Loomis and Knapp (2003), using a stereoscopic HMD system, investigated egocentric depth perception in a virtual environment. Virtual spheres were presented to observers on the ground plane at 2, 6, and 18m. A triangulation task was carried out where observers had to view the sphere, turn away, and proceed to walk with vision for 3m. At this point, observers were required to point back at the sphere; using two points of reference.

Assessing pointing errors showed observers underestimated distance by a factor of 2 in virtual compared to real environments (Loomis & Knapp, 2003). In a depth matching experiment Livingston et al (2003) tested observers' perception of virtual objects against that of real objects from distances of 5.3m to 44.2m. Placing 8 referents along a hallway, observers had to match either a real or virtual object with placed referents. Researchers were able to conclude that observers performed the task better with real objects compared to virtual. There has however been other work which has contrasted with the finding of compressed distance perception in virtual environments (Interrante et al 2004; Witmer and Sadowski 1998).

Comparing blindfolded walking in a real hallway to blindfolded walking in a virtual hallway, Witmer and Sadowski (1998) further analysed comparisons between real and virtual environments. In both conditions observers viewed a target and walked without the use of their vision. It was reported that mean errors of distance judgment ranged between 1 – 11% for real environments and between 2 – 18% in the virtual condition. Mean error increased with increasing target distance, but no significant difference between environments was reported. It was also found that observers made more errors in both

environments when tested in the virtual environment first. Interrante et al (2004) found similar capability for participants to walk without vision for upto 10m in both a real room and in a high-fidelity virtual room presented in a HMD.

The contradictions from findings are split between experiments using visual controlled designs and action based studies (see *table 1.5*). Table 1.5 presents a number of studies assessing depth perception in virtual environments, alongside the method and accuracy of depth judgments. The table highlights how depth perception differs hugely in virtual environments and could be largely dependent on the method used.

This suggests that in an AR environment how depth is measured is as crucial as in real-world. Experiments have produced similar trends as found in real-world environments; action based studies (Philbeck & Loomis, 1997) produced better depth estimates than studies implementing a visually controlled design (Loomis et al, 1992).

Despite these studies findings (Loomis & Knapp 2003; Livingston et al 2005; Witmer & Sadowski 1998; Interrante et al, 2004) of a varied outcomes of distance perception in virtual environments. There is a general consensus in the field that depth perception in virtual scenes is distorted compared to real-world scenes (Plumert, Kearney, Cremer, 2005).

Examining depth perception in AR Swan II et al (2007) measured medium and far distances in AR depth perception. Carrying out two experiments, first a perceptual matching task, the second a combination of verbal report and a blind walking task, depth judgments were measured at distances between 5.25 and 44.31m. The first experiment (perceptual matching task) observers used an optical see-through AR display.

Study	Distances (m)	Method	Accuracy
Witmer and Sadowski (1998)	4.6 – 32.	Treadmill Walking	85%
Knapp (1999)	5.0 – 15.0	Triangulated Walking	42%
Durgin et al. (2002)	2.0 – 8.0	Direct Walking	65%
Willemsen and Gooch (2002)	2.0 – 5.0	Direct Walking	81%
Willemsen et al. (2004)	4.0 – 8.0	Triangulated Walking	47%
Willemsen et al. (2004)	4.0 – 8.0	Direct Walking	62%
Thompson et al. (2004)	5.0 – 15.0	Triangulated Walking	44%
Interrante et al. (2006)	3.0 – 9.2	Direct Walking	95%
Jones et al (2011)	3.0 – 7.0	Direct Walking	90%

Table 1.5: A table displaying previous studies findings (in terms of distance accuracy), distances tested and method used for virtual reality environments (source: from Jones et al 2008, Table 2 (p.14)).

A series of eight differently coloured, real-world referents were equally positioned down the hallway (See *Figure 1.17*). The virtual target was a semi-transparent rectangle which horizontally filled the hallway, and vertically occupied half the hallway's height. Observers were required to adjust the virtual target's depth using a trackball, until it matched the depth of the referent in the hallway of the same colour.

Results showed a significant result of increased error with increased distance, with an underestimation bias found below 23m, which switches to overestimation of depth at distances above 23m (Swan II et al, 2007). This finding was consistent with the weakened strength of stereo disparity through

medium-field depth. A finding which is similar to real world depth perception (Palmisano et al, 2010).

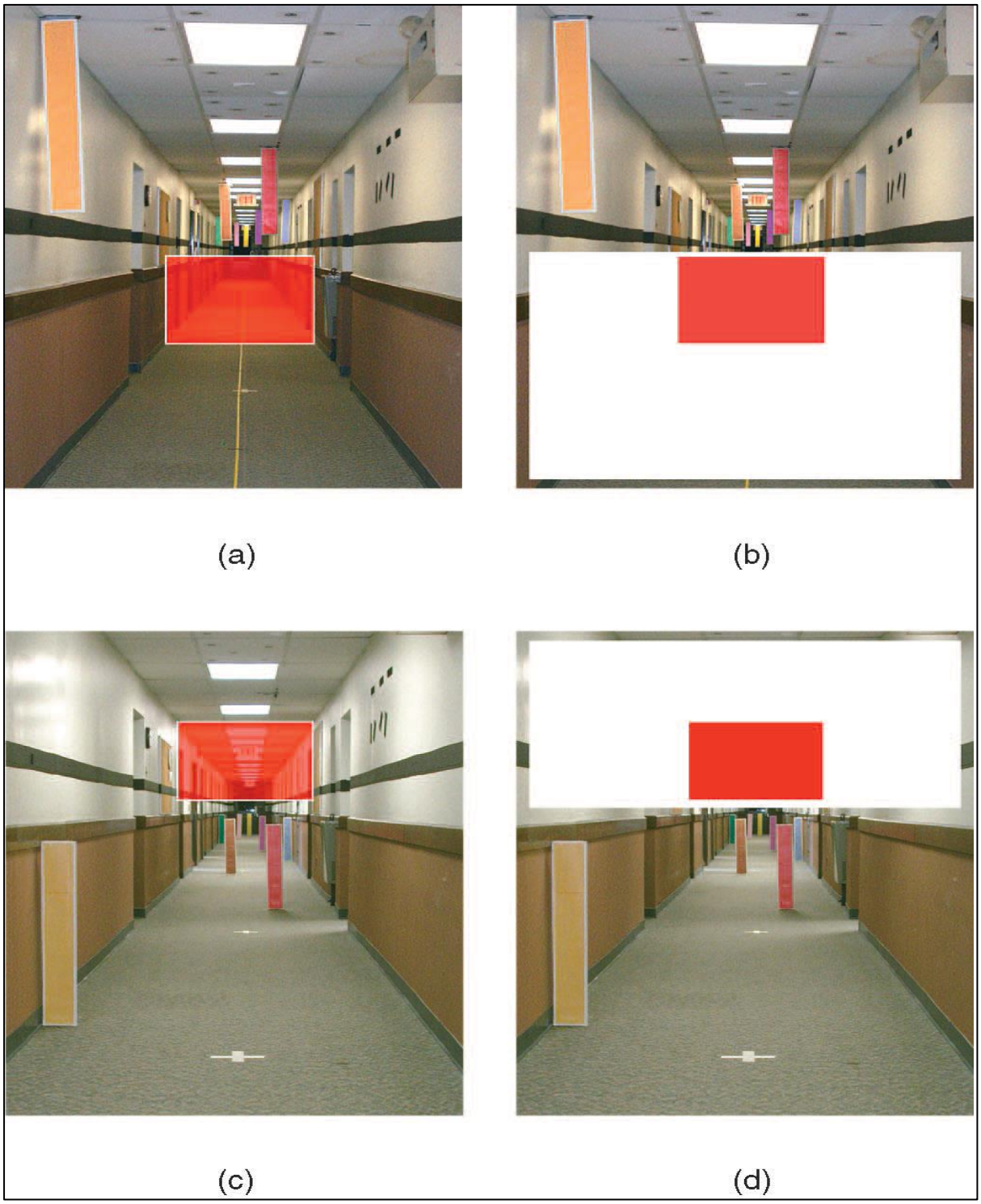


Figure 1.17: The experimental setting and layout of the real-world referent and the virtual target rectangle for Jones et al (2008). The images are not photographs taken through an actual AR display, but rather are accurate

illustrations of what observers perceived (source: from Swan II et al 2007, Figure 1 (p.432)).

The second experiment, using a HMD applied combination of blind walking and verbal report tested observers between distances of 3m to 7m. The aim of the experiment was to characterise the depth underestimation phenomena in previous AR literature and to compare judgements of virtual objects and virtual objects that augment the appearance of real objects. As a control condition depth judgments viewed through an AR HMD display and real objects with an 'unencumbered view' were also examined. Observers' task was to judge the distance of a physical referent object (a 3D diamond) and a virtual model of the real-world referent object, in a blind walking procedure. Results again demonstrated the underestimation of egocentric depth perception for virtual objects presented in transparent AR HMDs.

Reported in terms of mean percentages of error (percentage = judged distance/actual distance) it was found that for Swan et al II (2007): 1) blind walking in the real world: 96%, 2) blind walking in HMD environment: 86% and 3) Verbal report 77**%.** Comparing these results, reported by Thompson et al (2004), the 86% reported error is at the high end which has been previously observed before (42% - 85%). When reporting in terms of mean error results from Swan II et al (2008) showed much more variability in judgements with verbal reports compared to blind walking procedure; further highlighting the variability in depth estimates in AR depth judgements due to methodology. The variable results represent a higher variance between methods.

Jones, Swan II, Singh, Kolstad and Ellis (2008) directly measured and compared egocentric depth judgments of AR and VR within the same experimental framework. Depth judgements were measured for 3, 5 and 7m,

again using the blind walking paradigm. Interestingly the results from the study revealed no underestimation of AR; an underestimation of VR depth was reported, although the effect was low compared to previous research.

The relatively small amount of error underestimation observed in the VR condition was contributed due to the device implemented in the study. The researchers claimed the HMD system used had the ability to be calibrated in 'AR mode' and then perform in 'VR mode. It was claimed this impacted on the results because calibration accuracy for a number elements (field of view, positon and orientation of trackers) could be set for AR but not VR conditions (Jones et al, 2008). It was concluded that the finding of no AR underestimation indicates that VR underestimation is due the real world background; as the inclusion of a real world background causes VR underestimation to disappear (Jones et al, 2008).

Findings from real-depth perception have established thresholds at long distance of 20m being 10%, and 40m being 20% of their actual distance (Palmisano et al, 2010). Overall depth perception findings in virtual environments have been inconsistent, with depth underestimation (where AR imagery is perceived to be further away than it is) results ranging from 6%, of actual distance compared to perceived distance, to as high as 68% (Kirkley, 2003). In addition some research has reported no difference between AR depth judgments and real-world depth judgments (Jones et al, 2008). These inconsistencies (see *Table 1.6)* could be due to the varying ways depth has been measured; such as direct walking, triangulated walking, and perceptual matching.

Study	Distances (m)	Method	Accuracy
Swan II et al. 2006	5.25 – 44.31	Direct Walking	82%
Swan II et al. 2007	3.0 – 7.0	Perceptual Matching	99%
Swan II et al. 2008	3.0 – 7.0	Direct Walking	79%
Jones et al (2011)	3.0 – 7.0	Direct Walking	96%

Table 1.6: A table displaying previous studies findings (in terms of distance accuracy), distances tested and method used for augmented reality environments

1.11.2 The use of HMDs

As well as the nature of the methodology used to examine depth, one thing the above studies have in common is that they all examined virtual distance perception using HMD systems. These devices allow for only a limited FOV, both in horizontal and vertical directions. As we know from real world depth perception, decreased FOV results in a compression of depth perception, with large FOV increasing perceptual gain (Todd, Thaler, Dijkstra, 2005). However research investigating the impact of limited FOV has produced similar estimates for real and virtual stimuli with increased FOV (Plumert et al, 2005) or shown that when observers FOV is restricted, to the extent experienced in HMD, no impairment for distance perception in the real environment is established (Loomis & Knapp, 2004).

Jones et al (2011) suggest the issue of restricted FOV and inability to see visual information in the periphery as the cause for the underestimation. This has been already indicated previously by Willemsen et al (2004); Creem-Regehr et al (2005); Wu et al (2004).

A vast range of different head-mounted display technologies have been used to test AR. For example, in AR environments an increase in contrast with the background will cause an area appear closer (O'Shea, Blackburn & Ono 1994). In optical see-through systems the contrast offset is even higher; with

Grechkin et al (2010) establishing that high contrast of non-photorealistic augmented objects caused a significant underestimation of the distance.

Over years of research AR underestimation has been accredited to limited FOV, monocular vs stereo viewing, the weight and inertia of the HMD, the effects of implicit feedback and the quality of the rendered graphics. Research findings have found that some of these factors do contribute to the underestimation in AR; however none of them have fully defined it. A range of experimental measures and a number of different devices to project virtual imagery are likely to account for the vague understanding in AR depth perception. In the thesis a series of experiments to measure AR depth perception in a novel method; without the use of a head-mounted display, in a wide open setting, and controlling for depth cues. Results from the experiments will give clearer answer to the underestimation effect in AR, how does AR and real world depth perception differ, and present a model for depth perception in AR.

1.12 Summary

The development of AR technology and its utilisation in large display devices, in terms of both field of view and depth, allows for better understanding of current psychological theories and concepts. As well as developing practical outcomes for devices such as vehicle HUDs.

In terms of field of view, wide displays present imagery further in a user's periphery, possibly whilst simultaneously carrying out a primary central task. Previous research examining the effects UFOV has demonstrated two types of findings; tunnel vision and general interference. It has been suggested that findings are a result of the specific experimental design implemented. For

example task with a sufficiently difficult foveal task will result in tunnel vision (Williams, 1988, 1989).

This is similar to previous work on attentional theory detailing early and late attentional selection, where for example stimulus parameters such as presentation time of stimulus and demand of foveal task resulted in either early or late attentional selection (Sperling, 1960; Eriksen & Eriksen, 1977; Kahneman & Treisman 1984). Tunnel vision/general interference and early/late selection map together closely as the demand on one task directly impacts on the ability to complete the other.

Also few studies have also examined field of view at eccentricities above 90^{o} (Van Wijgert, 2003) with most studies in both UFOV rarely examining attention larger than 60^{o} display size (Mackworth 1960, Ikeda & Takeuchi, 1975; Leibowitz & Appelle, 1969; Holmes et al 1977; Chan & Courtney, 1998), and early/late attention above 30^{o} (Sperling, 1960; Eriksen & Eriksen, 1977; Kahneman & Treisman 1984).

Applying perceptual load theory predictions on early and late attention selection Chapter 2 (comprising of two experiments) will examine participants UFOV using a demanding foveal task whilst manipulating both difficulties of foveal and peripheral tasks. Experiments will manipulate presentation time and peripheral target eccentricity for presentation upto 120^{o}.

In terms of depth perception previous work in AR has been varied, with reports of underestimation and depth error ranging vastly (Kirkly, 2003). A possibility for this is because of the wide range of devices and methodologies used to carry out AR depth judgments (Renner, Velichkovsky, & Helmert, 2013). Also due to technological issues with the use of HMDs as the preferred presentation method for AR, it is difficult to match results with depth perception in the real-world. As real-world depth for long distance judgments has been

recorded implementing perceptual matching tasks (Allison et al 2009, Palmisano et al, 2011), AR has measured depth using action based walking paradigms. When comparing design methodologies in real-world environments depth judgments have been found to differ from study to study (see *table 1.5*).

Chapter 3 (comprising of six experiments) presents an experimental design for AR implementing design procedures based on real world experiments (Allison et al, 2009) assessing a variety of cues to depth in AR. Results will help understand how accurate AR depth perception is and how present to present and highlight real world objects using virtual imagery.

Chapter 2: Field of view and attentional workload

Abstract

Devices using augmented reality have the potential to become much larger; for example AR HUDs which encompass an entire windscreen. This will require users to acquire information from their periphery, often whilst carrying out another task in their foveal vision (UFOV). Previous examinations on the human capabilities of UFOV and attentional resources have only tested on limited display size. Research which has tested on large displays, such as driving related studies, has done so without applying some key attentional concepts and frameworks. To examine the interaction between attentional workload and useful field of view, two dual-task attention experiments were carried out. Participants were required to control a moving bar with a steering wheel (Critical Tracking Task) in their foveal vision, whilst completing a concurrent peripheral Landolt C orientation (250 msec per presentation) discrimination task across a horizontal plane (120^{o}). Fixation task difficulty was manipulated by the increased difficulty of the Critical Tracking Task (increased speed and decreased size), with peripheral Landolt C difficulty manipulated by the visibility and eccentricity of presented peripheral targets. In experiment 1 peripheral performance decreased as a function of eccentricity, demonstrating an effect of early attentional selection. In the follow up experiment the impact of a dual-task, increased vertical eccentricities and presentation time of peripheral Landolt C (1 second) targets was examined. Results revealed peripheral performance did not decrease as a function of eccentricity, with results demonstrating late attentional selection. It was concluded UFOV is time-dependent rather than eccentricity-dependent, with design implications for AR HUDs described.

The chapter is split into two sections; the first focuses on peripheral workload for target eccentricities increased across a horizontal plane with a short presentation time (250 msec). The second section focusses on the effect of dual-task processing, increased vertical eccentricities and increased peripheral target presentation time (1 s).

Section A will be begin by describing potential attentional issues derived from the use of large AR devices and related driving research examining UFOV. Following predictions under the framework of perceptual load theory will be described. The chapter will then describe the method and design for the first experiment examining the interaction between foveal load and peripheral load, detailing the critical tracking task and peripheral Landolt C task. Following results and discussion of findings in terms of attentional theory and practical design implications for AR HUDs will be presented.

Section B will follow with first discussing the effect of increased vertical presentation eccentricity in UFOV and peripheral presentation time with regards to previous research examining early and late attentional selection. The second experiment of the chapter will be described, followed by the results and a discussion of the findings in terms of early and late attentional selection and design implications for AR HUDs.

SECTION A: Workload and horizontal eccentricities

2.1 Introduction

AR is being utilised for applications which demand both high levels of attention (e.g. medical and driving applications) and the ability to split attention between the AR system and the surrounding world. For example in a vehicle HUD a driver has to be able to follow directions presented by virtual imagery whilst simultaneously having to detect real-world road markings to keep in lane.

Current AR technology is often applied in devices small in size and perceivable FOV; such as tablets, smart phones and current vehicle HUDs (around 20cm x 30cm). Reviewing early research into the use of small HUDs in both aviation and vehicle applications Gish & Staplin (1995) claimed that small HUDs are only advantageous in low attentional workload conditions. After reviewing research on HUDs, Gish & Staplin (1995) concluded the small available FOV demands too much of the users attentional resources culminating in attentional tunnelling. For example with a small vehicle HUD attentional tunnelling could manifest as a driver focusing resources on the small presentation of imagery in front of them and potentially ignoring potential road side hazards, road turnings, pedestrians. Would attentional tunnelling still occur with a larger HUD?

In 'Section A' attentional issues raised by potential AR devices with a wide FOV will be defined, followed by an examination of driving simulator research examining foveal and peripheral tasks. This will lead to an experiment which examines both increased foveal and peripheral load in a psychophysical simulation. The fixation task is represented by a critical tracking task, which simulates the specific driving task of in-lane keeping, with the peripheral task a Landolt C discrimination task. Following results of the testing, findings will be discussed in terms of early and late selection within the attentional framework

of perceptual load theory and practical implications for AR HUDs with wide FOV.

2.1.1 Attentional Tunnelling

Attentional tunnelling occurs under circumstances where attention is allocated to a certain channel of information or task for a duration which has a detrimental impact on other channels; resulting in being unable to perform other tasks (Wickens & Alexander, 2009). For example a peripheral visual task performance being hampered by the presence of a foveal task. Previously described tunnel vision (early selection) and general interference (late selection) can be understood in the frame work of attentional tunnelling. When attention is allocated to a certain channel, for instance a foveal channel, peripheral performance either decreases as a function eccentricity (tunnel vision/early selection) or across all eccentricities (general interference/late selection).

2.1.2 Perceptual load theory predictions

Perceptual load theory (Lavie 1995; Lavie & Tsal, 1994 – see 1.6.1) proposes that the filtering out of visual distractors in our periphery will likely depend on the 'perceptual load' of a current attentional task at the fixation point. When the central fixation task demands attentional capacity, less processing of the peripheral field (thus less interference from distractors) occurs. Therefore as attentional load for a central fixation target increases there is a greater potential for the omission of irrelevant peripheral targets at an earlier stage of processing.

Although there has been a lot of support for the theory in laboratory settings (Lavie, 1995; Lavie & Fox, 2000; Beck & Lavie 2005), there has been a limited examination for targets presented far in the periphery. Applied work in driving research using driving simulators (Miura, 1990; Ball & Owsely, 1992; Crundall, Underwood & Chapman, 1999; Seya, Nakayashi & Yugi, 2013) has explored peripheral distractors at a wider FOV; however the task of driving has

sometimes been passive in nature (e.g. Crundall et al, 1999, 2001), findings may be applicable only in the specific environment (Seya et al, 2013), and the perceptual load of peripheral distractors has often not been considered.

For instance in driving full windscreen AR HUDs may present imagery to highlight potential hazards in the periphery which will require the users to respond quickly and correctly. It will be necessary for users to be able to perceive such imagery, in both low and high load driving conditions.

On other occasions AR imagery will highlight points of interest such as road turning, street names, and buildings. In these instances imagery can be presented for longer as the user will not be expected to make an instant driving response. However it is important in these circumstances peripheral imagery does not distract a user from the central driving task.

Early work on studies investigating UFOV using a dual task paradigm found the effect of tunnel vision; increase in foveal task load resulted in a decrease in peripheral performance. These studies link with experiments assessing early/late selection hypothesis; the experiments mentioned above used simple foveal tasks (i.e. letter identification) with displays upto 90^{o} in size. Applying findings of tunnel vision to potential full windscreen HUDs, when visual imagery is presented in to the far periphery, users are likely not to perceive the information. This is because the foveal task (e.g. keeping control of the vehicle) will demand too much of a drivers attentional resources.

Thus how exactly does attentional selection occur for dual tasks with a demanding cognitive fixation task, and peripheral targets presented far in the periphery of UFOV. Does early or late attentional selection occur? It is

important and necessary to evaluate AR devices in terms of early and late selection as it will implicate how designers present imagery on large devices.

The current chapter presents a dual-task attention study presenting participants with a fixation task and peripheral vision task. The fixation task implements a critical tracking task (CTT) as a measure of fixation performance. Based on the design of Lansdowne, Brook-Carter & Kersloot (2002), the CTT requires a users' continuous attention. The task involves participant controlling a virtual bar using a steering wheel; the task is to keep the bar within the boundaries of another larger moving bar (see *Figure 2.1).* There will be three different conditions for the CTT, representing difficulty; easy, medium and hard difficulties. Task difficulty will be manipulated by the speed and size of the outer boundary. The CTT was implemented as it has previously been implemented in similar studies testing In-vehicle systems (Petzoldt, 2011) and it puts a form of demand on the participant that is fairly similar to various aspects of driving (Groeger, 2000). The performance measure for the CTT was number of lane exceedance. Each time the participant crossed the boundary of the outer black box.

Peripheral vision was assessed by a Landolt C orientation discrimination task. Landolt C target were presented in one of four orientations (see Figure 2.2). Targets were presented in one of four contrast conditions and were also displayed in one of six locations across the same horizontal plane; 60^{o}, 40^{o}, 20^{o} left and right of the participant.

2.1.3 Research Question & Hypothesis

Can participants accurately discriminate targets far in their periphery for short presentation times when fixation task requires additional attentional demands? How do findings relate to theories of attentional processing and load?

- Hypothesis 1: Increase of fixation difficulty will affect target discrimination
    - Prediction 1.1: Peripheral target discrimination performance will deteriorate as a function of eccentricity
        - Tested by ANOVA to determine a significant interaction between fixation difficulty and peripheral target position

- Hypothesis 2: Increase of peripheral task difficulty will not affect fixation task
    - Prediction 2.1: Fixation task will not produce increased errors as peripheral task difficulty increases
        - Tested by ANOVA, a non-significant interaction between fixation load and peripheral load

2.2 Method

2.2.1 Participants

Thirty participants took part in the experiment (mean age = 22.2 years, SD = 3.11), 24 were undergraduate or postgraduate students of the University of Nottingham, and 6 members of the general public. There was no requirement for participants to possess a driving licence as testing focussed on the limitations of basic visual perception. Demographic and driving information was recorded for each participant. Power analysis

Participants were recruited by using advertising posters at University of Nottingham campus and through a recruitment system for undergraduate psychology students, as a requirement for their undergraduate degree course.

2.2.2 Stimulus & Apparatus

The experiment was presented on a large curved screen used typically for driving simulation experiment. The images were presented on the screen using three projectors. The size of the projection was 60 x 192cm. The visual angle was calculated using measurements of car windscreens and research findings which found the 95^{th} percentile measurement distance from the windscreen to the driver eyes. Participants used a PC connected 'Thrustmaster Ferrari GT Experience' steering wheel to control and respond to stimuli. Four buttons on the right hand side of the steering wheel corresponded to the four different orientations of the Landolt C stimuli. The steering wheel was used to control the central white box moving left or right.

The size of the Landolt C target were 18^{o} in size. The CTT consisted of a center white box; an outer black box and an outer white boundary (see *Figure 2.1).* The outer white boundary was a size of 20cm x 192cm, the center white box dimensions were 7.5cm x 15cm. The size of each Landolt C target was 18^{o}.

The CTT consisted of an inner white box (8^{o} x 4^{o}) controlled by the participant and an outer black box (12^{o} x 6^{o}) which consistently moved from left to right within the confines of an outer white boundary (100^{o} x 6^{o}). The entire display was 120^{o} x 30^{o}.

The outer black box took a variable time to cover the distance it moved, taking 1 to 1.5 seconds to travel from one side to the other side of the range. The distance range covered by the black box was 16^{o} with the outer black box moving in the opposite direction once it reached its set boundary position on either the left or right side of the participant's visual field, with a frequency of cycling ranging from 0.5Hz – 0.7Hz. The participants task was to keep the inner white box inside the outer box, thus the movement of the outer box was

constructed in a manner which would not be too predictable. Pilot work was carried out two participants to determine the parameters of the CTT

As the outer black box moved left and right, with the participant controlling the inner white box to keep in the confines of the black box, the white box moved nearer and further from the Landolt C's presented in the periphery. Participants were tracking the white central box, therefore their eyes moved to track the box. The range of eccentricities for peripheral Landolt C targets was 10^{o} – 70^{o} from the observer.

The size and speed of the black box was manipulated for increased difficulty (see *table 2.*1). The experiment was created using Psychopy v1.74.

CTT	Size of black box	Speed (Hz)
Easy	**12^{o}**	**0.5**
Medium	**10^{o}**	**0.5**
Hard	**10^{o}**	**0.7**

Table 2.1: Critical tracking task difficulty manipulations in terms of size and speed

Landolt C Condition	Luminance (cd/m^2)
0.2	**41**
0.3	**62**
0.5	**103**
0.7	**144**

Table 2.2: Peripheral Landolt C target contrast and luminance (cd/m^2)

2.2.3 Design

The experiment implemented a dual task within-subjects design involving a peripheral target discrimination task and a critical tracking task. The independent variable of the CTT was difficulty, which was manipulated by increasing the speed and decreasing the size of the CTT outer boundary box, into three levels (easy, medium and hard).

The dependent variable was number of lane exceedance. Each time the participant crossed the boundary of the outer blue box. The independent variables for the peripheral target discrimination task were the contrast ratio, set at four different levels (0.7, 0.5, 0.3, and 0.2 – see *Table 2.2)* and six horizontal positions on the large display (60^o, 40^o, 20^o left and right of the participant – see *Figure 2.1*). There were four possible orientations of the Landolt C (see *Figure 2.2).* The dependent variables for the peripheral task were the ability to correctly identify the orientation of Landolt C target, and participant reaction times.

Trials lasted between 2 – 3 minutes. A Landolt C would appear on average every 2 – 2.5 seconds. Each Landolt C was presented for 0.25 (s) and a total of 80 were presented each trial in four different orientations.

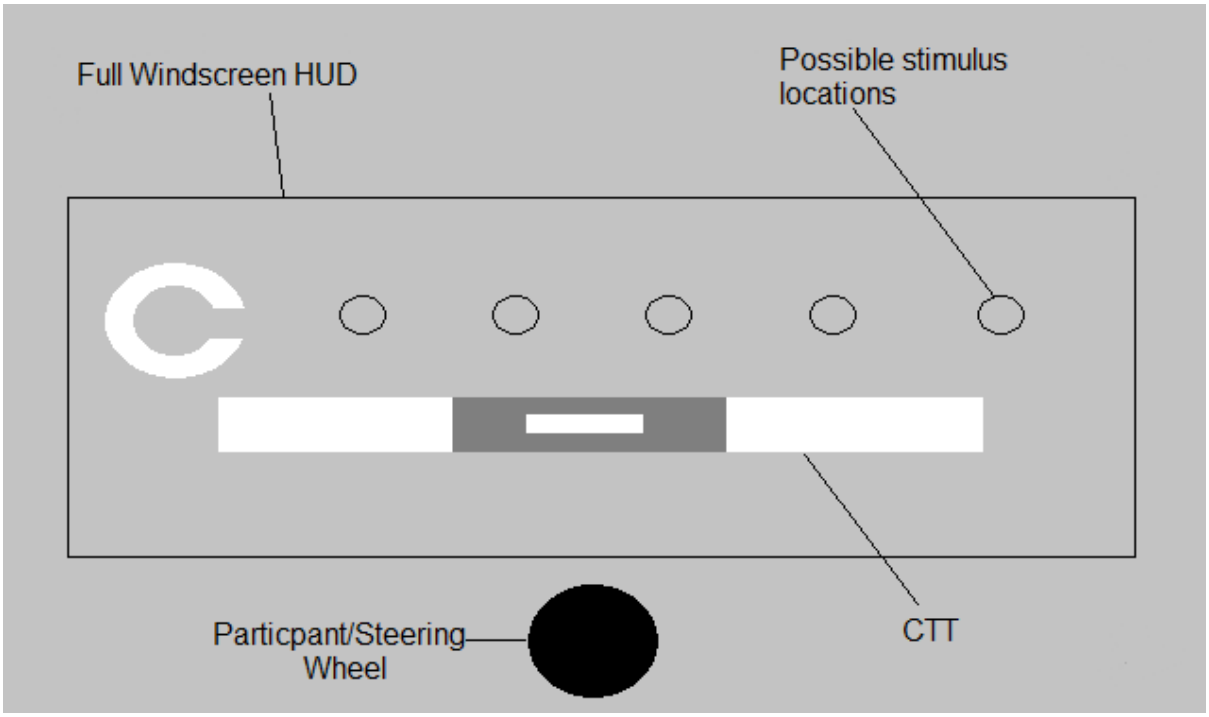


Figure 2.1: The arrangement of peripheral Landolt C targets in relation to the critical tracking task

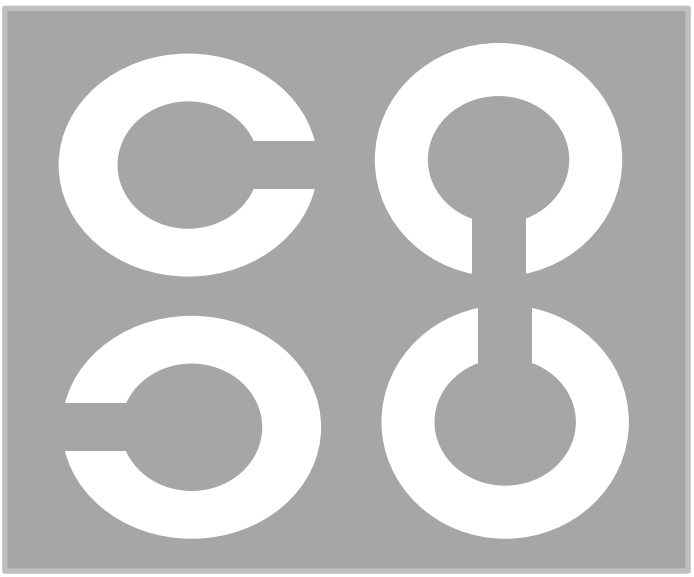


Figure 2.2: Schematic representing the four possible orientations of Landolt C targets.

2.2.4 Procedure

The experiment was conducted in NITES 2 driving simulator at the University of Nottingham. The experiment used the large display of the simulator however the vehicle rig was removed and replaced with a small table and steering wheel. Under reduced room illumination participants were seated at the steering wheel. To become familiar with the combination of CTT and the peripheral discrimination task participants first had one practice run through the CTT, followed by a practice trial discriminating the orientation of the Landolt C targets. A final practice trial involved participants being accustomed to completing both task simultaneously. After participants had practiced and were sufficiently comfortable with the setup the experiment began.

Participants were instructed that the primary task was the CTT, and only attend to peripheral targets if they can. In the main experiment, with participant sat at the steering wheel, the display would first present the CTT. Participants would be controlling the center white bar, as the outer black bar with be moving left and right. Whilst trying to keep the white bar within the boundaries of the black bar, after 2 to 2.5 seconds a Landolt C target would be presented at one of the six peripheral locations, in one of four possible orientations for 0.25 seconds. Participants were required to respond to Landolt C using one of four corresponding buttons on the steering wheel. The experiment would last until 80 peripheral Landolt C's had been responded to.

The experiment involved a total of 24 trials, each trial taking between 2 – 3 minutes. Participants would complete eight trials at a time, taking up to 20 minutes and then have a five minute break. They would complete this sequence two more times. Trials were counterbalanced to avoid order effects. Each participant took part in all trials.

2.3 Results

The following results section has been split into the two hypotheses stated in the introduction:

- Hypothesis 1: Increase of fixation difficulty will affect target discrimination
- Hypothesis 2: Increase of peripheral task difficulty will not affect fixation task

Performance in the fixation task (CTT) was measured by number of errors. The performance of the peripheral task (Landolt C) was analysed by detection rates (performance on being able to detect target) and orientation discrimination rates (ability to correctly identify the correct orientation of the Landolt C).

2.3.1 The impact of foveal load on peripheral performance

2.3.1.1 Peripheral discrimination performance

To assess the impact of increased fixation difficulty on a peripheral orientation discrimination task, a three-way 3 (fixation difficulty) X 4 (peripheral target contrast) X 6 (peripheral target position) within-subjects ANOVA was carried out on participants peripheral discrimination scores. After assumptions of sphericity were met, a significant three-way interaction was established (F (30,870) = 1.7, p = .011, η_p^2 = .056), as well as a main effect of peripheral target position (F (5, 145) = 119.8, p < .001 η_p^2 = .805). Post hoc tests using a Bonferroni correction found that increased eccentricity results in a drop off in performance. The mean difference from presentations presented at 20^o against 60^o presentations on the left was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions. The mean difference from presentations presented at 20^o against 60^o presentations

on the right was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions from the fixation task.

To assess the first hypothesis a two-way 3 (fixation difficulty) X 6 (peripheral target position) within-subjects ANOVA was carried out. After assumptions of sphericity were met a significant interaction was established (F (10,290) = 4.7, p < .0001 η_p^2 = .139) as well as a main effect of peripheral target eccentricity (F (5,145) = 120.9, p < .0001 η_p^2 = .806). Post hoc tests using the Bonferroni correction pairwise comparisons for peripheral target eccentricity found that increased eccentricity results in a drop off in performance. Paired sample t-test analysis compared each peripheral target position (20^o, 40^o, and 60^o) across fixation difficulty. Significant differences of fixation difficulty were only established at 60^o between easy > hard conditions (t=3.3, p = .002) on the right visual field. The interaction between foveal difficulty and peripheral target eccentricity potentially demonstrates the effect of early attentional selection, with post hoc analysis indicating at the furthest target eccentricity increased central workload has the impacts the most; with a potential shrinking of UFOV.

Further post hoc analysis examined the specific nature of the interaction to establish if peripheral target performance decreased as a function of eccentricity. A total of 18 paired sample t-tests were carried out. To correct for multiple comparisons the familywise error rate was calculated the alpha level was changed. The corrected critical value for 18 paired sample t-tests was set to p < 0.0028 (0.05/18 = 0.0028).

At easy fixation difficulty paired sample t-tests demonstrated significant differences between 60^o < 40^o (t=7.0, p < .001), 40^o < 20^o (t=4.1 p <.001), and 60^o < 20^o (t=9.4, p < .001) for presentations on the left. For peripheral target presentations on the right significant differences were found; 60^o < 40^o (t=5.2, p < .001), 40^o < 20^o (t=2.4 p =.017), and 60^o < 20^o (t=6.7, p < .001) with

decreased discrimination performance in the direction of increased target position eccentricity (see *Figure 2.3*).

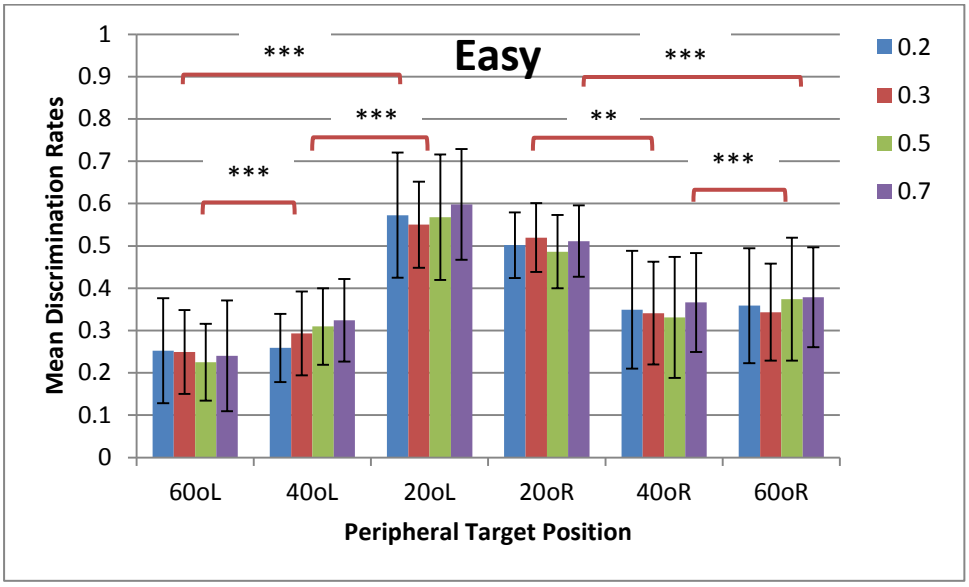


Figure 2.3: Participant peripheral target correct discrimination rates split by target contrast for easy, fixation difficulty conditions. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * (*p*<.05), ** (*p*<.01) and *** (*p*<.001).

For medium fixation difficulty paired sample t-tests demonstrated significant differences between $60^o < 40^o$ (t=4.4, $p < .001$), $40^o < 20^o$ (t=5.1 p <.001), and $60^o < 20^o$ (t=8.1, $p < .001$) for presentations on the left. For peripheral target presentations on the right significant differences were established; $60^o < 40^o$ (t=5.0, $p < .001$), and $60^o < 20^o$ (t=5.9, $p < .001$) with decreased discrimination performance in direction of increased target position eccentricity found (see *Figure 2.4*). No significant differences was found between $40^o < 20^o$ presentation positions.

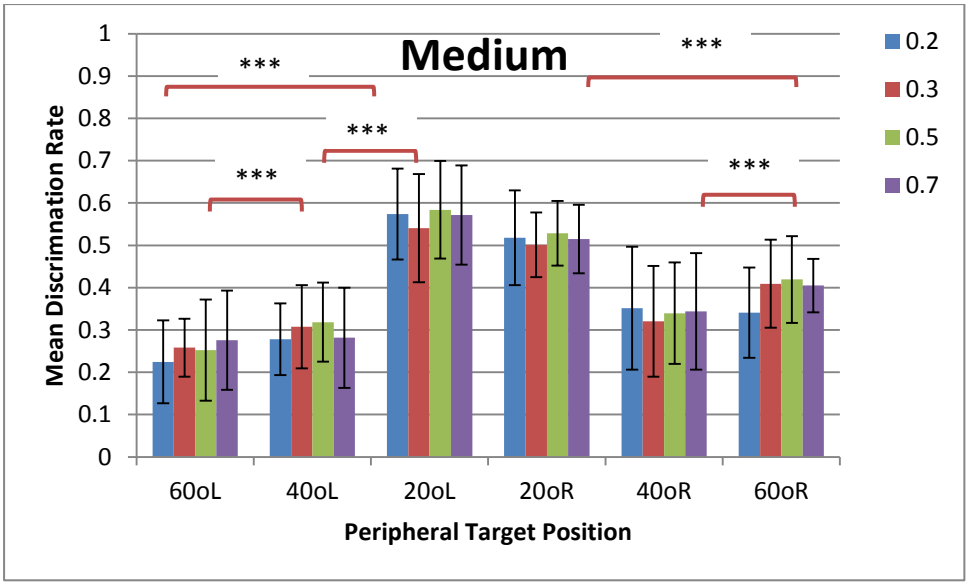


Figure 2.4: Participant peripheral target correct discrimination rates split by target contrast for medium, fixation difficulty conditions. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * (*p<.05), ** (p<.01) and *** (p<.001).*

For hard fixation difficulty paired sample t-tests demonstrated significant differences between $40^o < 20^o$ (t=3.9 p <.001), and $60^o < 20^o$ (t=5.1, p < .001) for presentations on the left visual field. For peripheral target presentations on the right significant differences were found; $60^o < 40^o$ (t=3.4, p = .001), $40^o < 20^o$ (t=2.5 p = .012), and $60^o < 20^o$ (t=5.2, p < .001) with decreased detection performance in direction of increased target position eccentricity found (see *Figure 2.5*).

Another set of post hoc comparisons, using paired sample t-tests, compared the differences between target positions at each workload to determine whether there was a decrease in target discrimination performance with increased task difficulty at fixation point. A total of 18 paired sample t-tests were carried out, with a Bonferroni corrected alpha level (p < .0028). Only one paired sample t-test produced a significant difference between fixation load;

60° right hand side between easy (.39) and hard (.32) conditions ($t(29)=5.4$, p <.0001).

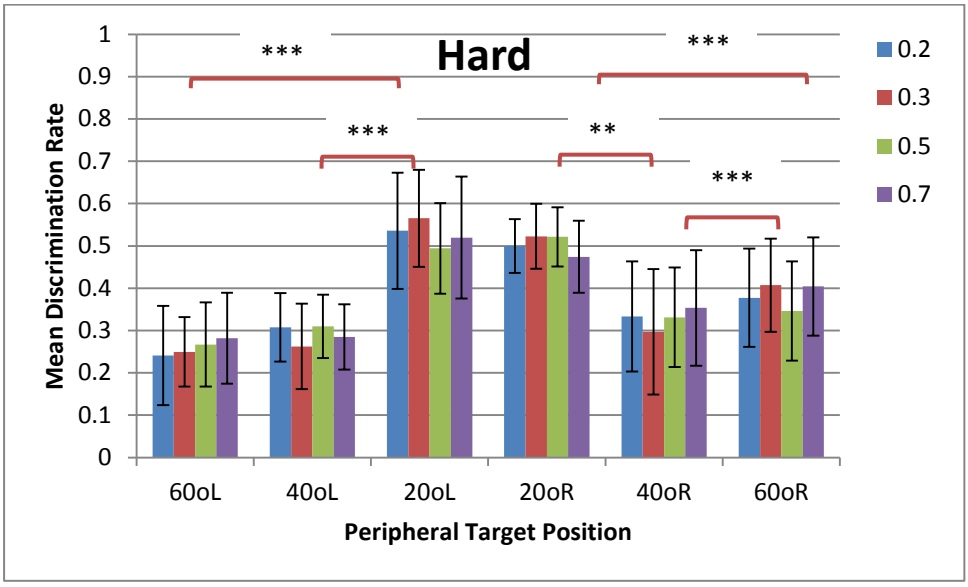


Figure 2.5: Participant peripheral target correct discrimination rates split by target contrast for hard fixation difficulty conditions. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * ($p<.05$), ** ($p<.01$) and *** ($p<.001$).

To explore each fixation condition, a set of two-way 4 (peripheral target position) X 6 (peripheral target contrast) within subject ANOVAs for easy, medium and hard were carried out. For easy workload the Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(119) = .50$, p <.0001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .53$). The result revealed no significant interaction, and a main effect of peripheral target position was established (F (5,145) = 78.4, $p < .0001$ $\eta_p^2 = .730$).

Applying a Bonferroni correction, pairwise analysis revealed that increased eccentricity results in a drop off in performance. The mean difference from presentations presented at 20° against 60° presentations on the left was established (p <0.001), indicating that peripheral performance decreased at

increased eccentricity positions. The mean difference from presentations presented at 20^o against 60^o presentations on the right was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions from the fixation task

For medium workload the Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(119) = .001$, p <.0001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .58). The result revealed no significant interaction. A main effect of peripheral target position was established (F (5,145) = 78.4, p < .0001 $\eta_p^2 = .730$).

Applying a Bonferroni correction, post hoc pairwise comparisons revealed that increased eccentricity results in a drop off in performance. The mean difference from presentations presented at 20^o against 60^o presentations on the left was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions. The mean difference from presentations presented at 20^o against 60^o presentations on the right was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions from the fixation task

For hard workload the Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(119) = .121$, p <.0001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .48). The result revealed a significant interaction of F (15,435) = 2.4 p = .021 $\eta_p^2 = .76$. The interaction indicates that at hard difficulty of the fixation task peripheral target conditions of eccentricity and contrast have more of an impact. A main effect of peripheral target position was established (F (5,145) = 78.4, p < .0001 $\eta_p^2 = .730$).Post hoc analysis applying a Bonferroni correction, pairwise comparisons revealed that increased eccentricity results in a drop off in performance. The mean difference from presentations presented at 20^o against

60^o presentations on the left was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions. The mean difference from presentations presented at 20^o against 60^o presentations on the right was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions from the fixation task.

2.3.1.2 Reaction times for peripheral targets

Further analysis on participants' reaction times to correctly discriminate the orientation of a peripheral stimulus was carried out. To assess the interaction between fixation difficulty and peripheral load a three-way 3 (fixation difficulty) X 4 (peripheral target contrast ratio) X 6 (peripheral target position) within subjects ANOVA was carried out. Assumptions of sphericity were met and a significant three-way interaction between fixation difficulty X peripheral target contrast X peripheral target position (see *figure 2.6*) was established (F (30, 870) = 1.6, p < .018, η_p^2 = .053). A main effect of peripheral target position (F (5, 145) = 6.8, p < .0001 η_p^2 = .189) and fixation difficulty (F (2, 58) = 3.5, p =.038 η_p^2 = .107). Applying a Bonferroni correction, pairwise comparisons revealed that the peripheral targets farthest from the participant (60^o) were significantly faster at responding than those at 40^o in the periphery. This pattern happened both on the left visual side (p = .019) and on the right visual side of the observer. Despite a main effect of fixation load, Bonferroni correction pairwise analysis did not supply significant differences between the conditions.

To further explore the interaction between fixation difficulty and peripheral visual performance analysis was split into three fixation difficulty conditions. A set of two-way 4 (contrast) X 6 (position) within-subject ANOVA were carried out. After assumptions of sphericity were met a significant interaction between peripheral target contrast and peripheral target position was found in only the easy workload (F(15, 435) = 1.9, p < .019 η_p^2 = .063). A main

effect of peripheral target position was found in all three conditions (see *Table 2.3*). No main effect of target contrast was established

Condition	F	p	η_p^2
Easy	3.0	<.014	.093
Medium	5.1	<.0001	.149
Hard	6.3	<.0001	.177

Table 2.3: Main effects of peripheral target position for all three fixation difficulty conditions

Comparisons using paired sample t-tests compared reaction time responses between peripheral target positions at each fixation difficulty. A total of 18 paired sample t-tests were carried out. The alpha level was corrected for multiple comparisons using the familywise error rate applying Bonferroni correction. This was calculated and the corrected critical value for 18 paired sample t-tests was set to $p < 0.0028$. Analysis provided five significant results. For presentations in the left visual field, significant differences established at 40^o easy to hard (p <.001), and medium to hard (p <.001). On the right hand side 20^o easy to medium (p <.001) and 60^o easy to hard (p <.001) and medium to hard (p =.002).

Three further paired sample t-tests (60^o left vs 60^o right, 40^o left vs 40^o right, 20^o left vs 20^o right) established significant differences for targets presented on the 60^o left being significantly slower than those presented at the 60^o right (t = 2.3, p = 0.021). Further analysis compared presentations at 20^o to 60^o, in both left and right directions was carried out. Paired sample t-tests found significantly faster detection responses at the 60^o presentation to the 20^o presentations in the left visual field (t =3.5, p <.001). These results demonstrate a speed accuracy trade-off. A speed accuracy trade off represent instances where there is a compromise between the demands for maximum accuracy (of the Landolt C task) and minimum reaction times (Wickelgren, 1977). Relating

back to discrimination data participants produced significantly faster responses for performance to peripheral Landolt C targets but at a cost with significantly inferior orientation discrimination performance.

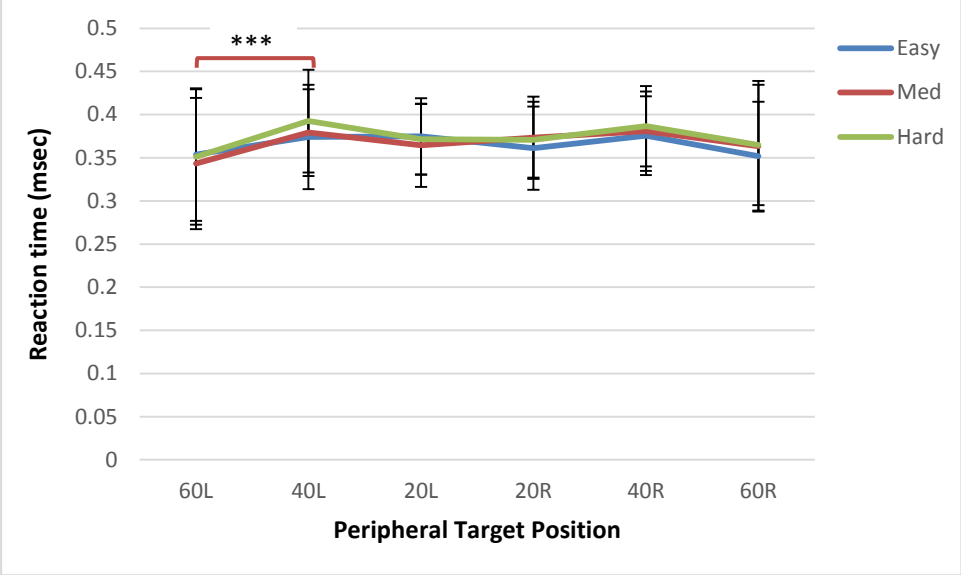


Figure 2.6: Participant peripheral target reaction times rates split by target contrast for Easy, Medium and Hard fixation difficulty conditions. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * ($p<.05$), ** ($p<.01$) and *** ($p<.001$).

2.3.2 Critical Tracking Task

The second hypothesis of the experiment examined the impact of increased perceptual load in peripheral vision on cognitive control for a foveal presented task; predicting participants should be able to protect the central task (Leibowitz & Appelle, 1969; Lavie 1995).

To assess the hypothesis measurements of errors on the fixation task were examined against decreased levels of peripheral target visibility. A two-way 3 (fixation difficulty) X 4 (peripheral target contrast) within-subjects ANOVA was carried out. The Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(20) = .010$, $p < .0001$, therefore degrees of freedom were

corrected using Greenhouse-Geisser estimates of sphericity (ε = .46). Results established no significant interaction (F (6,174) = 1.0, p = .42, η_p^2 = .033). A main effect of fixation difficulty was established (F (2,58) = 572.8, p < 0.0001, η_p^2 = .952).Post hoc test used Bonferroni corrected pairwise comparisons to reveal that easy fixation difficulty produced significantly fewer errors (see *Figure 2.7*) compared hard fixation condition (p <.001).

This finding supports the prediction proposed; an increase in perceptual difficulty would not impact on fixation task as participants were able to protect task-relevant and ignore task-irrelevant presentations in the periphery. One consideration with regards to the error data is does not take it to account participants time to recover for when they made the error to the time they manage to move back within the boundary.

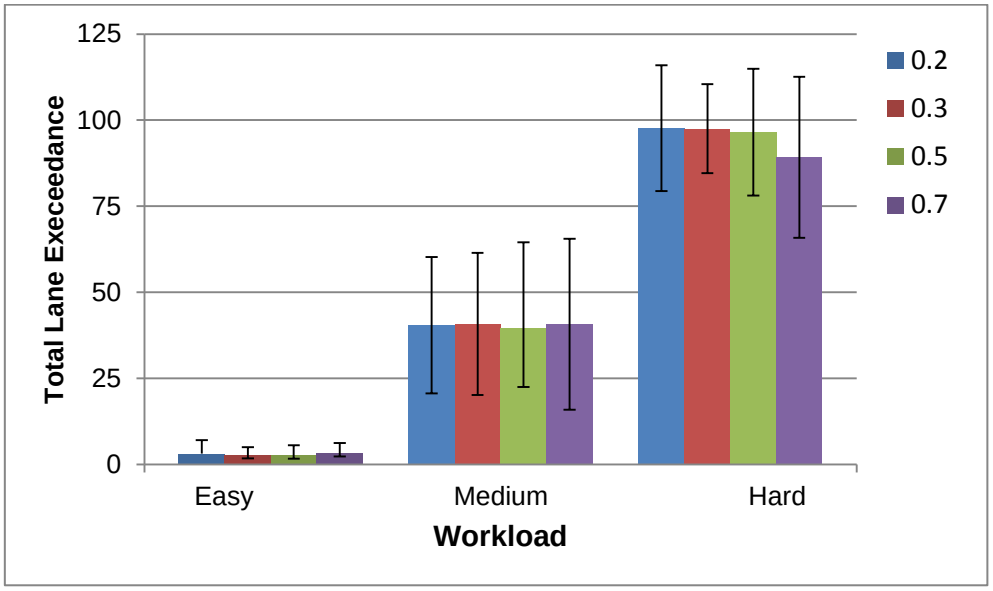


Figure 2.7: User performance as measured by average number lane exceedances in the easy, medium and hard fixation difficulty conditions. Peripheral target contrast levels tested at 0.2, 0.3, 0.5 and 0.7 level. Error bars represent S.E.M

2.4 Discussion

The first experiment of this experimental chapter investigated increasing levels of fixation load (difficulty on a driving-like task) and the impact it has on peripheral discrimination performance on a display representing the extremities of a vehicle windscreen (upto 120^{o} in a horizontal direction). The investigation was interested in exploring the interactions between fixation and peripheral load. Implementing a dual-task paradigm testing participants were required to control a moving bar in their central vision (critical tracking task) whilst simultaneously judging target orientation (Landolt C) in their periphery; splitting attentional resources between foveal and peripheral vision.

The aim of the study was to investigate the interaction between foveal and peripheral task loads further understanding the relationship between UFOV and eccentricity in terms of visual attentional theory. It is understood that environments which are complex in nature, that an area in space becomes more narrow, with a greater amount of attention required to process a greater amount of perceptual information

2.4.1 Cognitive load increases peripheral performance decreases

The first hypothesis explored the impact of increased cognitive load (presented in the fovea) on peripheral visual performance (presented on a simulated HUD). It was predicted that participants would protect the task-relevant stimuli in conditions of high fixation load with peripheral performance decreasing as a result; specifically as a function of target eccentricity (early selection).

Assessing participant's peripheral discrimination rates and reaction times, a significant interaction between fixation difficulty and peripheral target eccentricity was established; with an increase in fixation load reducing the effectiveness of peripheral vision as a function of eccentricity.

The result supports the prediction of early attentional selection. According to perceptual load theory perceptual processing becomes selective when capacity limits are reached (Lavie, 1995, 2005, 2010). When tasks demands surpass human capacity, this result in task-irrelevant items not being processed and ignored (early selection). Otherwise tasks which only enforce a small amount of demand, remaining attentional capacity are allocated to task-irrelevant items, causing distraction (late selection). Perceptual load theory also stipulates that humans' attentional capacity is impacted by cognitive mechanisms (e.g. working memory), and activities (e.g. driving a car) will determine the chance of distraction.

The results from the experiment demonstrates a foveal task requiring cognitive control demanded enough attentional capacity to cause peripheral task irrelevant items to be ignored when attentional capacity is reached for a large display.

The current experiment produced the finding of tunnel vision. Applying perceptual load theory, the finding represents early selection of attention. In the experiment a foveal task demanded a sufficient amount of attention which resulted in task-irrelevant items (peripheral targets) not being processed and ignored. Results from the current experiment suggest that early selection occurred as attentional capacity was potentially reached in all three fixation difficulty conditions, and therefore participants' perceptual capacity had been reached.

The interaction between fixation difficulty and peripheral target eccentricity supports past driving related research in driving simulators producing a similar result (Muira, 1990; Ball & Owesly, 1992; Crundall, Underwood & Chapman, 1999, 2002; Seya, Nakayshi & Yugi, 2013).

However the current findings do not support that of the recent experiment from Gasper (2016), producing an effect of late selection. Although the difference between the experiments revealing late selection, and the result of the current experiment, is the lack of a foveal task. Studies replicating an effect of early selection (e.g. Mackworth, 1965) tested participants with foveal tasks which demand limited attentional processing (e.g. letter identification).

The finding of early selection was potentially established for a variety of reasons. Firstly the experiment included a foveal task which was sufficiently difficult; unlike Gasper (2016). The experimental design also cohered to Williams (1988, 1989) requirements for producing early selection (or as described by Williams, tunnel vision); a foveal task which was given priority, and participants need to react fast to both foveal and peripheral tasks. The current experiment required participants to react quickly, as the surrounding box for fixation task moved fast, especially in the hard difficulty**.** Peripheral target presentations were only 0.5s, thus participants had to respond fast.

The current experiment implemented a dual-task paradigm; with previous conclusions arguing early selection in UFOV is only established in dual-task studies (Williams, 1988, 1989); with foveal and peripheral stimuli competing for attention, the foveal task was given priority (Leibowitz & Appelle, 1969). In the following experiment of this chapter a comparison between single task and dual task will be made.

2.4.2 Left visual field deficit

Analysis on the peripheral orientation discrimination task, interestingly found a significant difference in performance between left and right visual field presentations. In the current experiment both discrimination scores were found to be significantly inferior for left compared to right visual field presentations.

Participant response times were significantly slower on the left compared to the right. This is not the first time this specific finding has been established.

Seya et al (2013) also reported that RTs for presentations in the left visual field were significantly slower than in the right. In discrimination and visual search tasks performance is superior for horizontal directions compared to vertical directions (Carrasco, Talgar & Cameron, 2001; Carrasco, Williams, and Yeshurum, 2002; Corbett & Carrasco, 2011). Visual performance has also been shown to be improved in a lower direction compare to upper direction (Carrasco, et al 2001, 2002; Corbett & Carrasco, 2011). In the follow up experiment horizontal and vertical directions will be assessed and compared.

The finding raises the possibility UFOV on the left is narrower than on the right visual field. Accounting for similar findings, Seya et al (2013) applied the anticipation hypothesis which proposes that a driver's visual search is reliant upon anticipation of other vehicles; which is dependent upon driving experience (Hills, 1980)

Despite the anticipation hypothesis being originally used to explain the narrow UFOV in the upper direction of the visual field, Seya et al (2013) applied the concept to horizontal differences across the visual field in driving. Seya et al (2013) tested on Japanese drivers, who drive on the left hand side of the road, therefore to maintain the view of the entire road, as well as oncoming traffic, drivers will be expected to spread UFOV on the right more so than on the left. In the current experiment a combination of UK and Malaysian born participants were used; both countries which drive on the left hand side of the road (the same as Japan).

Another potential explanation for the difference between left and right visual field presentation is due visual constraints of the eye (Seya et al, 2013). It has been found that there is a higher amount of ganglion and cone cells in the

nasal (right) direction from the fovea, compared to the temporal (left) direction (Curcio et al, 1987; Perry & Cowey, 1985).

2.4.3 Protection of CTT?

The second hypothesis of the experiment investigated the impact increased perceptual load (reduced visibility of Landolt C targets) for peripheral vision has on a foveal presented cognitive control task; representing a specific driving task. It was predicted that a decrease in visibility of peripheral targets (increased perceptual load) should have no significant impact on task–relevant CTT task; no report of an increase on driving errors.

Analysing the number of errors made on the central fixation, as perceptual load increased there was no increase in errors on the fixation task; supporting the hypothesis reported. The result demonstrates participant's ability to protect a central task over peripherally presented information. A finding that replicates previous results and conclusions (Lavie, 1995, Beck & Lavie, 2005; Cartwright-Finch & Lavie, 2007). Leibowitz & Appelle (1969) concluded that when both central and peripheral stimuli are competing for attention, precedence is given to central information processing. In the current experimental setup participants were able to protect the CTT (fixation task). Beck and Lavie (2005) concluded that as foveal vision has superior processing abilities compared to peripheral vision it is more difficult to ignore distractors in foveal vision.

The second hypothesis examined was a null hypothesis, that an increase in peripheral task difficulty will not affect the fixation task. There are problems in testing and accepting a null hypothesis. A null hypothesis shows no difference between two or more variables in a population. In the current context there was no difference in peripheral fixation loads, however caution must always be used when interpreting a null result as they come from underpowered

analysis, potentially only revealing a sampling error, as there could be a difference between conditions but it is not known due to the lack of statistical power.

2.4.4 Study limitations

The findings of early selection may have occurred due to the presentation speed of peripheral targets; only being 0.5s. This was done to minimise eye movements, and test the limits of human attentional resources. However peripheral target presentation time may have had a large impact on the results found; specifically the finding of early selection. It has previously been found in circumstances of longer peripheral target presentations results of general interference (Schieber & Benedetto, 1998).

In the context of large AR displays a variety of information will be displayed to a user, requiring varying amounts of duration time. For instance in a full windshield AR HUD, hazards (e.g. avoiding collisions) will flash up requiring drivers to respond quickly. However secondary information (e.g. navigation information) presented in the peripheral can be presented for longer before a driver needs to respond. Although, will increased peripheral information presentation time potentially cause distraction (late selection) and negatively impact on the central task? In the follow up experiment peripheral target presentation was increased from 0.5s to 2.5s.

The current experiment only tested on horizontal eccentricities and not vertical. It has been found that vertical performance is worse than horizontal (Carrasco, Penpeci-Talgor & Cameron, 2001). The current experiment can be criticised for only testing the horizontal eccentricities of a large AR device, however not in the vertical direction; therefore not a simulated full windscreen AR HUD. The follow up experiment tested participants both horizontal and vertical eccentricities.

2.4.5 Summary & Conclusion

The first experiment of the chapter examined participants' wide field of view for peripheral target eccentricities of upto 120^{o}, whilst concurrently carrying out a demanding foveal task. A significant interaction between foveal load and and peripheral load was established; with peripheral performance decreasing as a function of eccentricity. Applying the findings of the experiment under the framework of perceptual load theory (Lavie, 1995) it was concluded that early attentional selection occurred as the foveal task demanded sufficient amount of attention resulting in the task-irrelevant items (peripheral targets) not being processed and ignored. The finding demonstrated could potentially have occurred because of the presentation time of the peripheral stimuli (only 50msec), rather than the wide presentation position of the peripheral targets. In the follow up experiment a similar experimental paradigm was conducted, however peripheral targets were presented on two vertical planes and for an increased time (1s).

Section B: Increased presentation time and vertical eccentricities

2.5 Introduction

In the previous experiment, a dual-attention task investigated peripheral visual performance of 6 target positions across the same vertical eccentricity. A finding of early attentional selection was demonstrated; with increase in foveal load resulting in a decrease in peripheral performance. Previous findings have demonstrated peripheral vision is affected by dual-task workload (Williams, 1988, 1989).

2.5.1 Vertical target presentations

In the follow up experiment a similar design was used to examine the impact of peripheral targets for increased vertical, as well as horizontal eccentricities. It is understood visual performance is not consistent across the visual field at equivalent eccentricities (Carrasco, Talgar & Cameron, 2001), with superior performance being found on the horizontal compared to vertical meridian (Rovamo & Virsu, 1979; Rijsdijk, Kroon & Van der Wildt, 1980; Carrasco et al, 1995); an effect referred to as horizontal-vertical anisotropy. Results have also demonstrated superior visual performance in the lower compared to upper locations in the visual field (Smith & Edgar, 1990; Previc, 1990). It has been proposed that the performance difference between horizontal and vertical may represent ecological constraints with more relevant visual information more likely across the horizontal dimension of space (Carrasco et al, 2001).

Previous experiments assessing UFOV, where peripheral targets have been presented for the dimensions of a windscreen, have either presented on one vertical plane (e.g. Van de Weijgert, 2003) or have not directly compared presentations of different vertical plane. In the current section an experiment

will implement the same experimental paradigm as section A; however peripheral stimuli were presented on a vertical plane 10^{o} above fixation task.

2.5.2 Peripheral target presentation time

The outcome of early attentional selection in the previous experiment may have been a result of the short presentation time of peripheral stimuli. In a similar experimental setup to 'Section A' Schieber & Benedetto (1998) had participants focused on a central fixation to control eye gaze location, with a peripheral detection task of LEDs being presented upto 30^{o} to both left and right visual fields of participants. Every three to five seconds one of the peripheral targets would be illuminated; with durations set at 75, 100 and 125msec. Findings showed that with sufficient processing time (125msec) peripheral performance can be maintained well in the to the periphery (Schieber & Benedetto, 1998). However the fixation task was a passive task with participants not controlling the task, it was only used to keep the central gaze fixation of participants. The experiment display was 60^{o} in size; where as in the following experiment implemented a display size of 120^{o}.

The debate between early and late attentional selection has been categorised by differences in experimental designs (see *1.5.2*). In the previous experiment of the chapter a finding of early attentional selection was demonstrated. Peripheral targets were presented for 50msec. Previous demonstrations of early selection also presented peripheral target stimuli for short durations. For example Sterling (1960) study characterising early selection presented target stimuli for 50 msec. In testing perceptual load theory Lavie (1995) presented peripheral stimuli for 50msec and found early attentional selection. However in the original flanker task, typifying late selection, target stimuli was presented until a participant responded. Reaction times ranged between 400 – 500 msec (Erkisen and Eriksen, 1974).

In the current experiment, peripheral targets were presented for 1sec. Applying perceptual load theory it was predicted that late attentional selection will be demonstrated by participants.

2.5.3 Research Question & Hypothesis

Can participants accurately discriminate targets far in their periphery for increased presentation times when fixation task requires additional attentional demands? How do findings relate to theories of attentional processing and load?

- Hypothesis 1: Increase of fixation difficulty will affect target discrimination
    - Prediction 1.1: Peripheral target discrimination performance will not deteriorate as a function of eccentricity: late selection
        - Tested by ANOVA to determine a possible interaction between fixation task and peripheral eccentricity

- Hypothesis 2: Increase of peripheral task difficulty will not affect fixation task
    - Prediction 2.1: Fixation task will not produce increased errors as peripheral task difficulty increases
        - Tested by ANOVA, a non-significant interaction between fixation load and peripheral load

- Hypothesis 3: Peripheral performances will differ across horizontal visual fields
    - Prediction: left visual field presentations to be inferior to presentations on the right visual field

        - Tested by ANOVA and post hoc paired sample t-tests to determine significant difference between presentation fields.

- Hypothesis 4: Peripheral performances will differ across vertical visual fields
    - Prediction: upper visual field presentations inferior to presentations on the lower visual field
        - Tested by ANOVA and post hoc paired sample t-tests to determine significant difference between presentation fields.

2.6 Method

2.6.1Participants

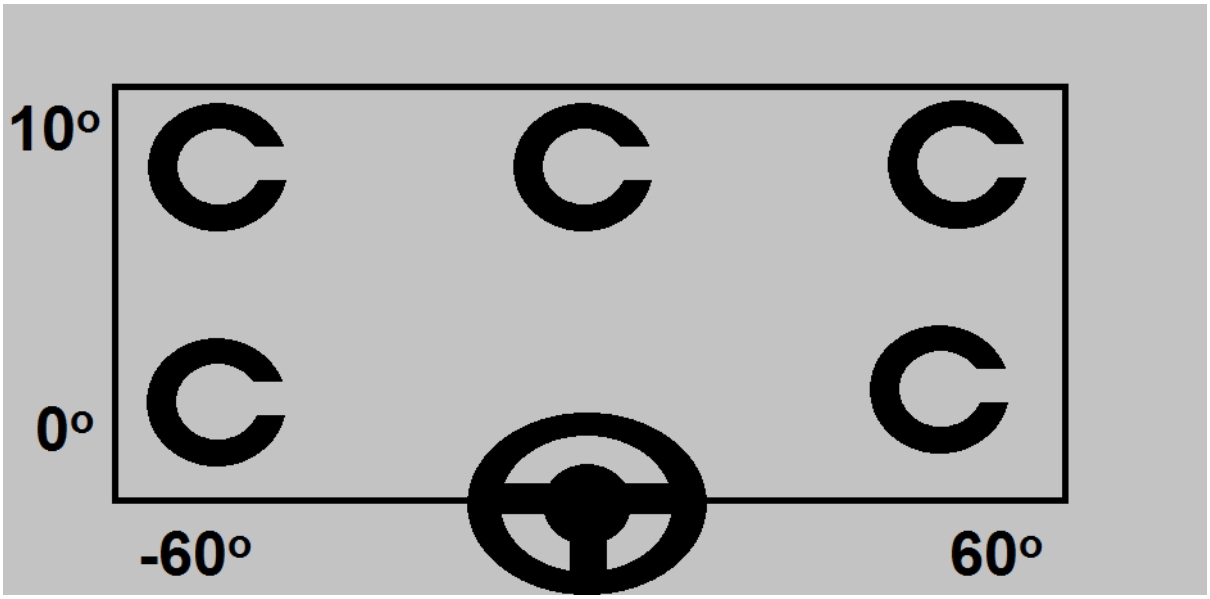
A total of 30 participants (18 female, 12 male) took part in the study, a mean age of 21.3 (SD = 4.22) with an age range of 19 – 28 years. All participants were students at the University of Nottingham and were required to have a driving licence. The average driving experience of participants was 3.2 years (range: 2 months – 10 years), driving on average 4335 miles per year (range: 100 – 18000 miles).

2.6.2 Design

Landolt C's were used as targets for the peripheral orientation discrimination task. Presented in four orientations, and at five different locations; two on a lower plane (2^o above CTT) presented at 60^o to left and right of the participant, and three on a higher plane (10^o above CTT) presented 60^o to the left and right, and one centrally at 0^o (see *Figure 2.8).* Size of peripheral targets was presented at 18^o. Landolt C's were presented for up to a 1s, and

presented on average every 3s. A minimum of 1.5s had to pass before another target could be presented, and a total of 55 targets were presented.

A within-subject design was implemented, the independent variables were target contrast (0.2, 0.3, 0.5 and 0.7) target position (5 levels, 2 lower plane, 3 higher plane), and CTT difficulty (easy, hard and single task). The effect of dual-task processing was tested by implementing a single task condition. The visual scene of the screen remained the same; however participants were asked to keep their gaze on the middle of the screen and to only use the four buttons to respond to the Landolt C targets, and not to move the steering wheel. In this condition the CTT was programmed so the white bar would never exceed the boundaries of the moving black bar. The dependent variable for the CTT were errors made (each time white box exceed boundaries of black box), and correct response rate and reactions times for peripheral orientation discrimination task.



F*igure 2.8*: Schematic representation of position of target presentations

2.6.3 Procedure

Participants were seated at distance of 80cm away from the screen. Participants were allowed time to get familiar with the combination of both the CTT and the detection task. First they were given five minutes on the CTT alone, and followed by five minutes alone on the peripheral detection task. A final practice trial (2-3 minutes) involved participants being accustomed to completing both task simultaneously. A total of 8 trials per CTT difficulties (four levels of contrast repeated twice) each trial would last up to taking up to 3 minutes. A five minute break was given to participants every 20 minutes, or at any time they requested. Trials were counterbalanced to avoid order effects. After the starting practice session, trials were randomly ordered for each participant.

2.7 Results

2.7.1 Increased foveal load on peripheral task performance

To investigate the interaction between foveal and peripheral task performance analysis participant's orientation discrimination performance for the peripheral targets was assessed. After assumptions of sphericity were met a three-way 3 (fixation task difficulty; no task, easy, hard) X 4 (peripheral target contrast: 0.2, 0.3, 0.5 and 0.7) X 5 (peripheral target position: lower left, lower right, upper left, upper center and upper right) within-subjects ANOVA established a significant three-way interaction, $F(24,672)=1.9$, $p = .004$ $\eta_p^2 = .065$). Main effects of peripheral target position ($F(4,112)=483.9$, $p < .0001$ $\eta_p^2 = .945$) and peripheral target contrast ($F(3,84)=63.2$, $p < .0001$, $\eta_p^2 = .693$) were found. Post hoc examination of the main effect of peripheral target position was done with Bonferroni corrected pairwise comparisons. Analysis revealed that presentations on the lower right were significantly superior to lower left ($p <$

.001). Upper right presentations were significantly superior to presentations on the upper left (p <.001). Lower left presentations compared to upper left were not significantly different, however lower right presentations were significantly superior to presentations on the upper left (p = .005).

Peripheral presentations in the top center (95%) were significantly superior to all peripheral target presentations apart from lower right (92%). The first hypothesis states that peripheral performance would not decrease as a function of eccentricity, this finding supports that prediction. The top center presentation position was 20^{o} away from the fixation center whereas the lower right was 60^{o}; and no significant difference was established between presentation positions (see *Figure 2.9*).

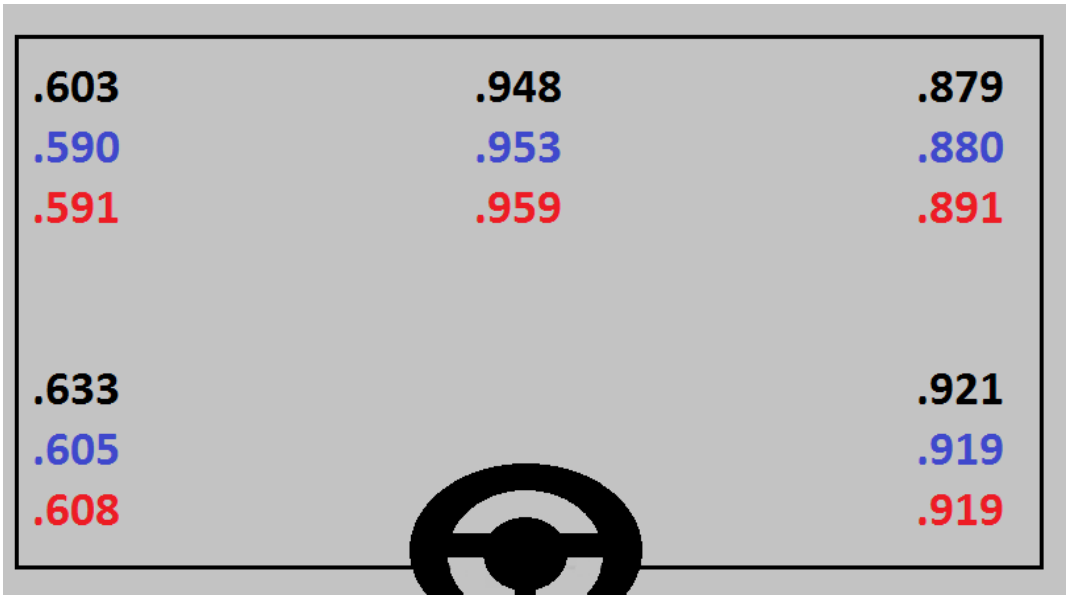


Figure 2.9: Schematic representation of overall discrimination score for each target for fixation difficulty: ST (black), easy (blue), hard (red)

With regards to hypothesis stated in the introduction (3 & 4), peripheral task performance was superior on the left side of the visual field compared to the right. Presentations lower on the visual field were only superior to upper field when presented on the right hand side. Post hoc analysis using Bonferroni corrected pairwise comparisons revealed as peripheral target visibility increases

orientation discrimination performance increases. Significant differences in performance were found between 0.2 and 0.3 contrast conditions ($p < .001$), and between 0.5 and 0.7 ($p = .028$). No significant result established between 0.3 and 0.5 contrast conditions (*see figure 2.9*). When comparing fixation conditions, or no fixation task, pairwise comparisons revealed no significant differences between conditions. Single task measure was found not to be significantly different to either easy ($p = 1.0$) or hard ($p = 1.0$) fixation task conditions. The same result found between easy and hard conditions ($p = 1.0$

To examine the nature of peripheral discriminations under various manipulations of fixation load a two-way 2 X (fixation difficulty) X 5 (peripheral target position) within-subjects ANOVA was carried out. Assumptions of sphericity had been met and no significant interaction between foveal load and peripheral target eccentricity (F (8,920) =.94, $p = .479$, $\eta_p^2 = .008$). The finding demonstrates the effect of late attentional selection, because as foveal difficulty increased peripheral performance did not decrease as a function of eccentricity.

A final set of analysis exploring the three-way interaction examined fixation difficulty and peripheral target position across peripheral target contrasts. After assumptions of sphericity were achieved a 2 (fixation difficulty) X 5 (peripheral target position) within-subject ANOVA across peripheral target contrast 0.2 condition established no two-way interaction. However a main effect of peripheral target position was found (F (4,112) = 194.2, $p <.0001$ $\eta_p^2 = .874$). To compare the impact of increased eccentricity on discrimination, mean differences for central (upper center) to peripheral targets on the left (UL) and right (UR). Bonferroni corrected pairwise comparisons revealed significant differences for presentations in both left ($p<0.0001$) and right ($p = 0.04$) directions of increased eccentricity.

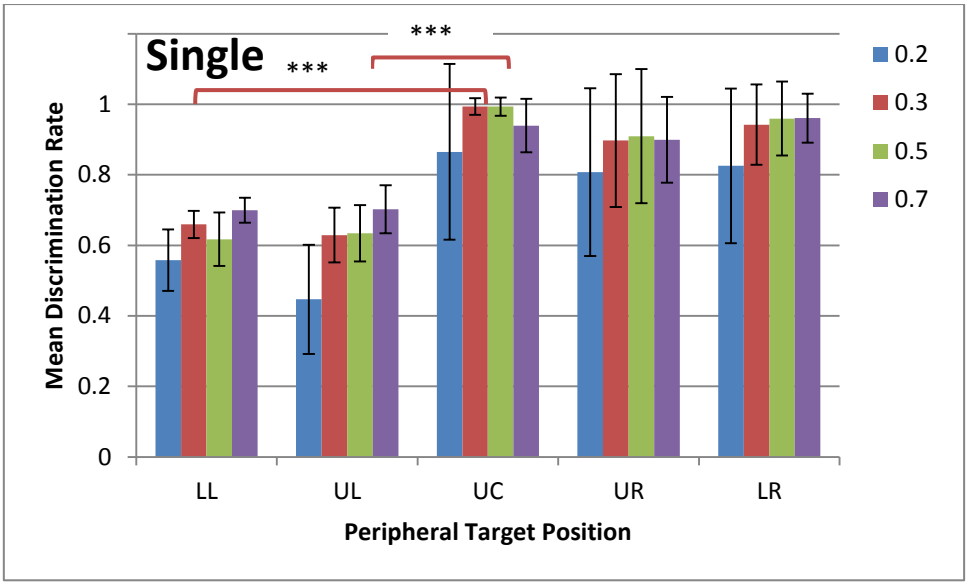


Figure 2.10: Participant peripheral target correct discrimination rates split by target contrast for single task fixation difficulty conditions. LL represents lower left position, UL represents upper left, UC represents upper center, UR for upper right, and LR for lower right. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * (*p<.05),* ** *(p<.01) and* *** *(p<.001).*

A 3 X 5 within-subject ANOVA across peripheral target contrast 0.3 met assumptions of sphericity but did not establish two-way interaction, however a main effect of peripheral target position (F (4,112) = 201.699, p <.0001 η_p^2 = .878) and fixation difficulty (F (2, 56) = 201.699, p =.004 η_p^2 = .183) were both discovered. To compare the impact of increased eccentricity on discrimination, mean differences for central (upper center) to peripheral targets on the left (upper left) and right (upper right). Bonferroni corrected pairwise comparisons revealed significant differences for presentations in both left (p <0.0001) and right (p = 0.001) directions of increased eccentricity. To compare the condition of fixation difficulty a Bonferroni corrected pairwise comparisons revealed

between single task and easy (p = 0.02) and single task and hard fixation task conditions (p = 0.007) were found.

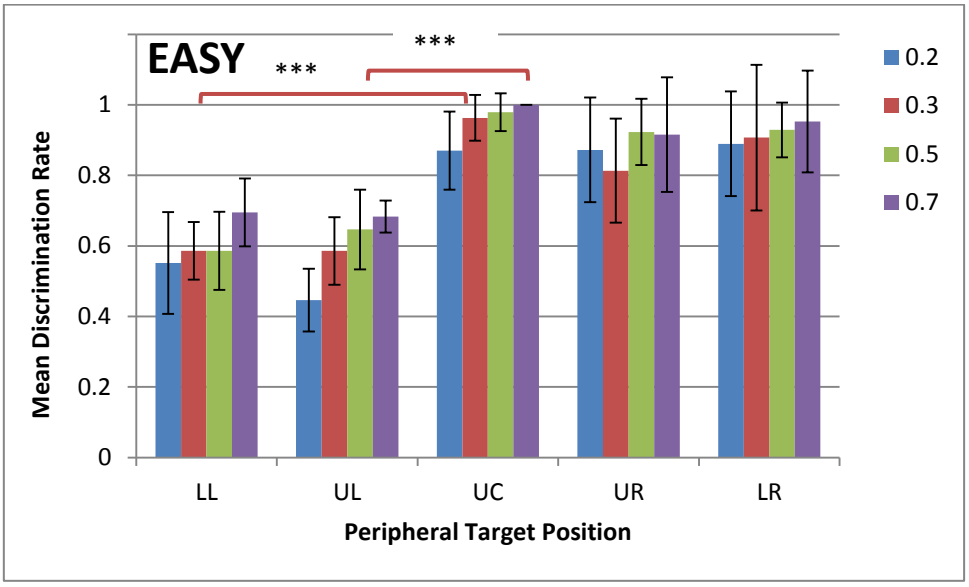


Figure 2.11: Participant peripheral target correct discrimination rates split by target contrast for easy fixation difficulty conditions. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * ($p<.05$), ** ($p<.01$) and *** ($p<.001$).

For contrast condition of 0.5 another 3 X 6within-subject ANOVA was carried out. With assumptions of sphericity being met no two-way interaction was established, a main effect of peripheral target position was found (F (4,112) = 194.2, p <.0001 η_p^2 = .897.) To compare the impact of increased eccentricity on discrimination, mean differences for central (upper center) to peripheral targets on the left (UL) and right (UR). Bonferroni corrected pairwise comparisons revealed significant differences for presentations in both left (p<0.0001) and right (p = 0.001) directions of increased eccentricity.

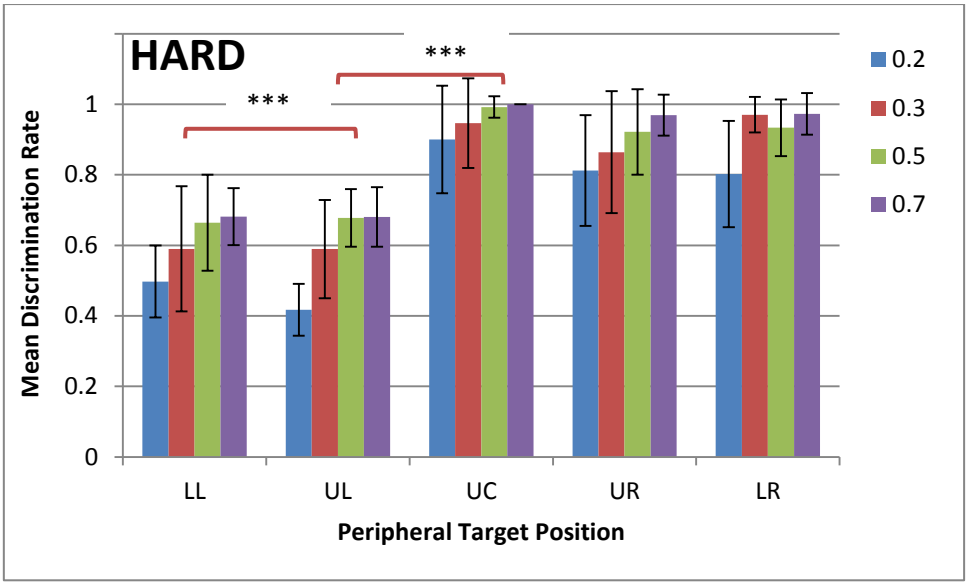


Figure 2.12: Participant peripheral target correct discrimination rates split by target contrast for hard fixation difficulty conditions. Errors bars represent S.E.M. Significant differences between peripheral target conditions represented * (*p<.05), ** (p<.01) and *** (p<.001).*

A final two-way 3 X 6 within-subject ANOVA across peripheral target contrast 0.7 was carried out. The Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(35) = .016$, p <.0001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .51). A significant interaction was established (F (8,224) = 3.5, p = .001 η_p^2 = .110), a main effect of peripheral target position was also found (F (5,145) = 219.2, p <.0001 η_p^2 = .887). The mean difference from presentations presented at 20^o against 60^o presentations on the left was established (p <0.001), indicating that peripheral performance decreased at increased eccentricity positions. Bonferroni corrected pairwise corrections revealed the mean difference from presentations presented at 20^o against 60^o presentations on the right was p <0.001, indicating that peripheral performance decreased at increased eccentricity positions from the fixation task.

The interaction between fixation difficulty and peripheral target position at the two most visible contrast conditions indicates that increased eccentricity of peripheral targets has more of an impact at higher workloads when peripheral targets are more visible.

2.7.2 Increased peripheral load on foveal task performance

Using errors made on the fixation task data, analysis applied examined how the visibility of the peripheral target affected task completion of the CTT. A two-way 3 (fixation difficulty) X 4 (peripheral target contrast) within-subjects ANOVA was carried out. Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(5) = .001$, $p < .0001$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .34$). Results established a main effect of fixation difficulty ($F(1,38) = 488.1$, $p < .0001$, $\eta_p^2 = .763$), however no significant interaction between CTT difficulty and peripheral target difficultly was found. This result suggests take that peripheral discrimination task did not affect performance on the driving task. Bonferroni corrected pairwise comparisons revealed significantly fewer errors ($p < .0001$) for easy compared to hard conditions (see *Figure 2.14*).

2.7.3 Reaction Times

Further analysis on participants' reaction times to detect a peripheral stimulus was carried out. To assess the potential impact of increased load of a driving task on peripheral detection reaction times a three-way 3 (foveal load) X 4 (peripheral target contrast ratio) X 5 (Peripheral target position) within-subjects ANOVA was carried out (see figure 2.13). After assumptions of sphericity were met no significant interaction was established. This further supports the argument for the presence of late attentional selection because the lack of interaction demonstrates that despite foveal difficulty increasing peripheral performance did not decrease as a function of eccentricity.

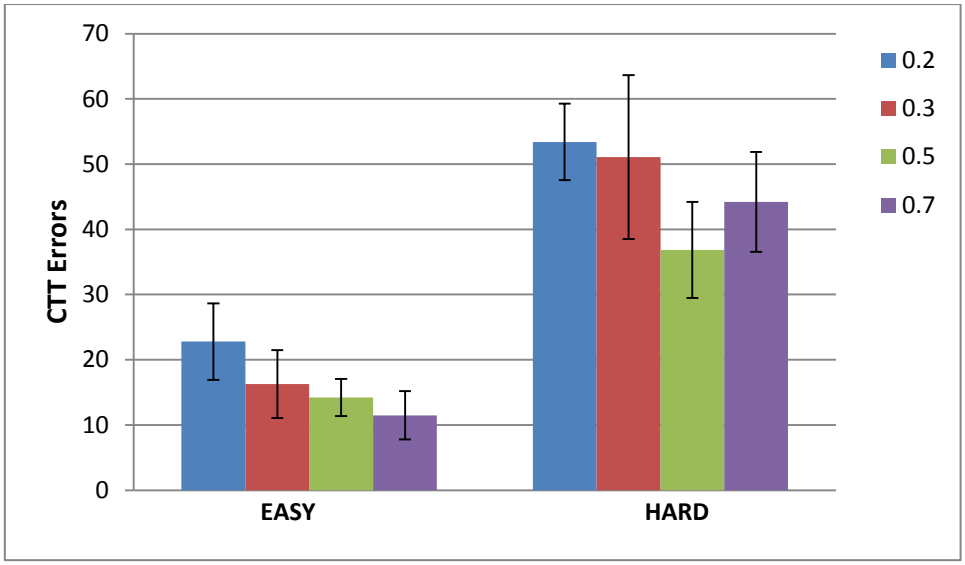


Figure 2.14: Driver performance as measured by average number of lane exceedances (Error bars = S.E.M)

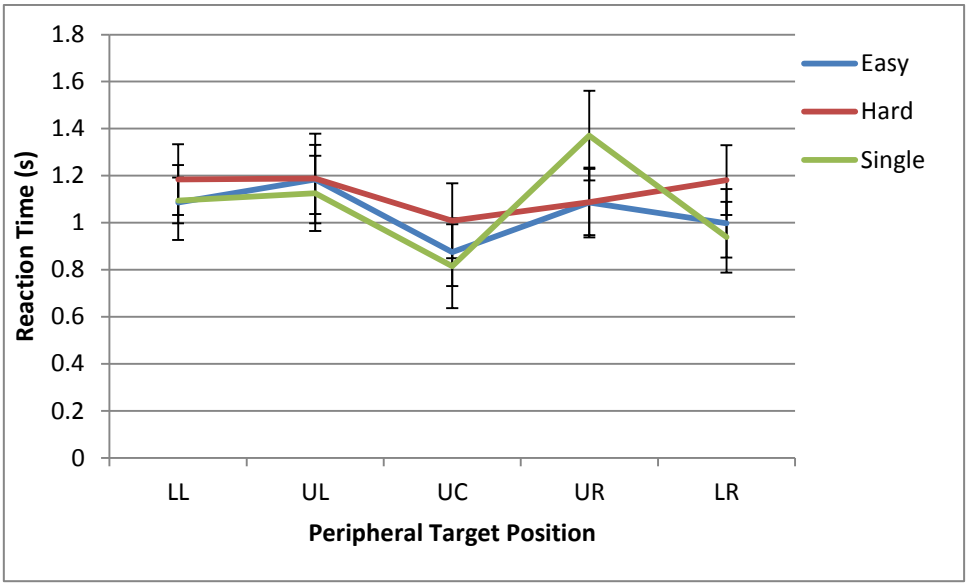


Figure 2.13: Participant peripheral target reaction times rates split by target contrast for Easy, Hard and single task difficulty conditions. Errors bars represent S.E.M

2.7.4 Peripheral target appearance

Further analyses examined the impact an appearance of a peripheral target presentation had on the participant making a subsequent error on the foveal task. The data was gained by measuring each time an error on the CTT was made and a Landolt C appeared within the preceding 1s compared to CTT errors where a Landolt C had not appeared.

A two-way 2 (fixation difficulty) X 4 (peripheral target contrast) within subjects ANOVA was carried out. After assumptions of sphericity were achieved no significant interaction between workload and contrast was found. A significant main effect of workload was established ($F(1,28) = 91.0$, $p < .001$, $\eta_p^2 = .223$). Bonferroni corrected pairwise comparisons found easy difficulty was less likely to be impacted by peripheral targets than the hard fixation condition ($p < .001$). Table 2.4 shows the statistical analysis for the ANOVAs analysing the relationship between foveal and peripheral tasks.

A paired-sample T-test was carried out between the likelihood of a peripheral target appearing in the same trial as a user making a lane exceedance error, compared the likelihood of there being no peripheral target appearing in the same trial as a lane exceedance error. A mean weighted lane exceedance score of 32.89 for when a target was present during a trial and an increased mean weighted lane exceedance score of 35.11 (see *Figure 2.14*). A significant difference of $t(28) = -2.548$ $p < .01$. This result demonstrates that the presence of a peripheral target did not cause an error on the foveal task. This further highlights participant's ability to allocate resources to the foveal task.

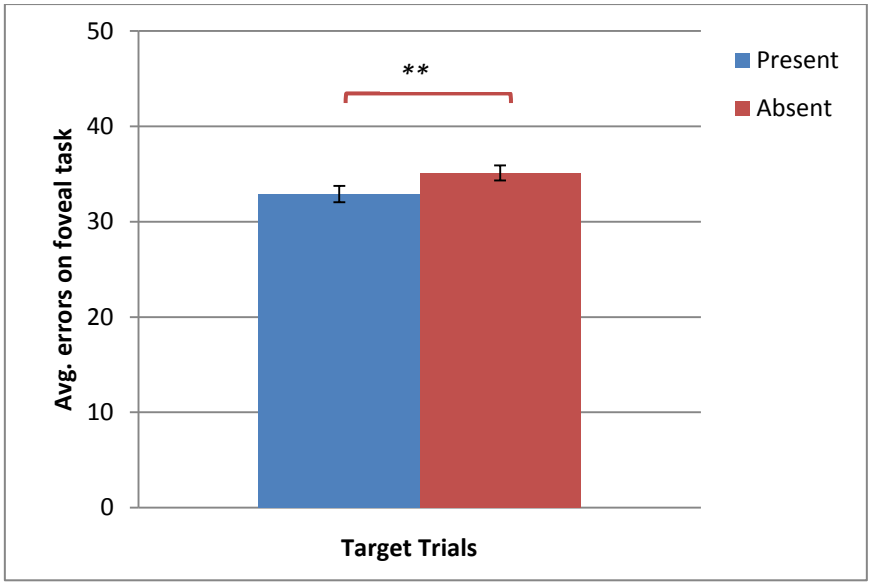


Figure 2.14: Graph representing the average number of errors on the foveal task for when a peripheral target was present or absent (Error bars = S.E.M)

2.7.5 Section A and Section B comparisons

In Section A when assessing the impact of increased difficulty of the fixation task on the impact of peripheral performance, an interaction was established; increased fixation difficulty resulted in peripheral performance decreasing as a function of eccentricity. Carrying out the same assessment for Section B found a different result. There was no interaction established with peripheral performance not decreasing as a function of late selection. Section A was interpreted as early attentional selection and Section B as late attentional selection.

To confirm that participant discrimination performance of Section A and Section B were significantly different to each other a three-way 2 (fixation task difficulty: easy, hard) X 4 (peripheral target contrast; 0.2, 0.3, 0.5, 0.7) x 2 (Experiment: Section A, Section B) between subjects mixed ANOVA was conducted. After assumptions of sphericity were met, analysis demonstrated a significant interaction ($F(3,969)=4.3$, $p=.005$, $\eta_p^2 = .013$) with main effects of peripheral target contrast ($F(3,969) = 66.2$, $p < .001$, $\eta_p^2 = .170$) and experiment

$F(1,323) = 622.3$, $p < .001$ $\eta_p^2 = .658$. Bonferroni corrected pairwise comparisons revealed Section A produced significantly inferior orientation discrimination results compared to Section B ($p < .0001$).

A limitation for the between subjects ANOVA is that it used contrast of the peripheral targets rather than peripheral target position. This was because peripheral target positions across the two experiments were different; however two peripheral target positions did match in Section A and B.

Comparing peripheral target positions at 60^o eccentricity for both left and right visual fields across fixation difficulty a two-way 2 (fixation difficulty: easy, hard) x 2 (peripheral position; left, right) x 2 (Experiment; Section A, Section B) between-subjects ANOVA was completed. After assumptions of sphericity were met no significant interaction was found. However, a main effect of peripheral target position ($F(1,234) = 789.7$, $p < .001$, $\eta_p^2 = .771$) and between subjects was established ($F(1,234)$, $p < .001$, $\eta_p^2 = 898$). Post hoc analysis used a set of four paired sample t-tests to compare discrimination results between experiment 1 and 2. Correcting for multiple comparisons the alpha level was set at $p < .0125$. All four comparisons revealed significant differences between experiment 1 and 2 (*see Table 2.4*).

Pair: Exp1 & Exp2	t-statistic	Sig
Easy Left	24.3	p<.001
Easy Right	33.8	p<.001
Hard Left	20.9	p<.001
Hard Right	35.5	p<.001

Table 2.4: Significant differences (using paired sample t-tests) for presentations at 60^o between experiments

2.8 Discussion

Following on from the initial finding of early attentional selection in the first experiment of the chapter, subsequent testing further explored the relationship between foveal load and peripheral performance. The current investigation increased vertical eccentricities of peripheral performance, increased target presentation time and examined the impact of dual task processing. The experiment implemented a similar experimental paradigm as the first experiment, testing examined participant's peripheral performance in terms of discrimination rates and reaction times. Foveal task performance was again measured in terms of errors. Findings from the experiment are interpreted in terms of early and late attentional selection, perceptual load theory and implications for large display sized AR devices (e.g. full windscreen vehicle HUDs).

2.8.1 Early or late attentional selection

The first hypothesis investigated the impact of increased central load on peripheral target eccentricity. The first experiment of the chapter established the effect of tunnel vision. In the follow up experiment peripheral targets were presented at increased vertical and horizontal eccentricities, however crucially were presented for a longer duration (from 0.25s to 1s). Analysing participant discrimination and reaction times to peripheral target presentations found no significant interaction between foveal load and peripheral target eccentricity; a result demonstrating the effect of general interference rather than tunnel vision. Even though in the second experiment peripheral targets were presented at increased eccentricities (in the vertical plane) compared to the first experiment, peripheral performance did not decease as a function of eccentricity.

The result supports other previous experimental work finding general interference (Holmes et al, 1977; Van de Weijgert, 2003; Gasper, 2016). The

main difference between experiment 1 and experiment 2, with regards to peripheral stimuli was the presentation time; increased from 0.5 – 2.5s. From this finding it is possible to support the conclusions that peripheral stimulus eccentricity may not be a factor in UFOV as long as participants have enough time process visual information (Seiple, Szlyk, Yang, & Holopigian, 1996; Schieber & Benedetto, 1998). Schieber & Bernedetto (1998) demonstrated peripheral visual performance is well maintained when processing time is sufficiently long. When available processing time is insufficient performance in peripheral vision suffers.

It is interesting that the same instructions were given in the first experiment as in the second; however presentation time of peripheral targets changed. It has been argued that instructions (Williams, 1988, 1989) maybe the most important factors with regards to the occurrence of tunnel vision or general interference; with priority given to the central task resulting in tunnel vision. However from the current findings peripheral stimulus presentation time may have a larger impact on the occurrence of tunnel vision or general interference.

Applying perceptual load theory (Lavie, 1995, 2005 & 2010), it can be suggested the current experimental results represent late-attentional selection, as remaining attention is not focussed on central task were allocated to the peripheral task-irrelevant items, causing distraction. The increased time peripheral target were presented for resulted in participants more likely to be distracted by them.

2.8.2 Fixation task priority

Analysing the number of errors on the fixation task found no significant impact of peripheral target contrast; demonstrating participant's ability to protect the foveal task, even with increased vertical eccentricities. This supports the hypothesis predicted in the introduction.

The finding demonstrates the precedence given to a central task over peripheral stimuli (Leibowitz & Appelle, 1969), with participants protecting the central task (Lavie, 1995; Beck & Lavie, 2005; Cartwright-Finch & Lavie, 2007).

In both Section A and Section B peripheral targets were randomly presented, in terms of presentation time, in occurrence to a completed in lane tracking from left to right. This factor of uncertainty of the exact presentation time of peripheral targets may have also prevented the deterioration of the foveal task (Van de Weijgert, 2003).

Another consideration is participants' instructions for the experiment; the foveal task was instructed to be the primary concern for participants, with the peripheral a secondary objective.

2.8.3 Horizontal and vertical eccentricities comparison

Analysis comparing presentations in the visual field for detection, discrimination and reaction times to peripheral targets found significant differences between left and right visual fields; a repeat result from the first experiment of the chapter. Results also revealed significantly inferior performance for upper visual field presentations compared to lower, however only on the left visual field.

The results demonstrated further support previous work identifying a narrowing of UFOV in the left visual field (Seya et al, 2013) and research findings demonstrating a horizontal superiority in peripheral performance compared to vertical directions (Carrasco et al, 2001; Carrasco et al, 2002; Corbett & Carrasco et al, 2011).

As there is more points of interest, especially in driving, across the horizontal meridian compared to the vertical it is not surprising that results often report a deficit in vertical plane perception compared to horizontal.

2.8.4 Implications for AR design with large display size

The finding of late selection, for instances of increased presentation time for peripheral targets, has potentially interesting ramifications for AR design of large devices. Designers need to be aware of peripheral information which is presented for too long on a screen, for information peripheral to the task of driving when presented need to reduce potential distraction.

This contradicts to the findings of experiment 1, of early selection. To compensate the occurrence of early attentional selection large AR devices could potentially utilise imagery which actives bottom-up activation of humans, for instance salient visual information in the far periphery if information is vital for the user. If a finding was late selection then guidelines when presenting information on large displays would encourage removing potentially distraction visual information in the periphery.

2.8.5 Summary and Conclusion

The first experimental chapter thesis explored the relationship between UFOV and attentional workload; with findings discussed in terms of perceptual load theory (Lavie, 1995, 2005, 2010) and the design of future AR devices. Two experiments were carried out, requiring participants to control a central task (representing the specific driving task of in-lane keeping) whilst completing a concurrent peripheral Landolt C orientation task on a large display (120^{o}); presented on the same horizontal plane. In the first experiment a significant interaction between foveal load and peripheral target eccentricity was established; revealing an effect of tunnel vision.

In the follow up experiment peripheral targets were presented across two vertical planes and presentation time was increased from 0.5s to 2.5s. Analysis reported no significant interaction between foveal load and peripheral target eccentricity, demonstrating a finding of general interference. The combination of results may indicate that peripheral target eccentricity is not a factor on UFOV when carrying out a concurrent foveal task; however peripheral target presentation time seems to be a more crucial variable.

It is concluded attentional selection in UFOV is time-dependent rather than eccentricity-dependent. When translated to the design of large AR device utilising increased UFOV it is proposed for visual information presented for user to react quickly, there is a need for it be presented in a bottom-up salient manner to avoid information being ignored. For visual information presented for longer durations, it is important it is designed in a manner which avoids distracting a user.

Over the course of two experiments, the chapter presented a study which is first to examine the UFOV comparing both horizontal and vertical increases in eccentricity whilst manipulating presentation time of peripheral targets. Also findings and conclusions were applied to perceptual load theory being the first to explain tunnel vision and general interference under the framework of early and late selection.

Chapter 3: Depth perception in augmented reality

Abstract

Virtual imagery in an AR scene can be presented at varying depths to highlight real-world targets/points of interest for a user. Past research and theories have suggested that AR depth perception differs significantly compared to VR and real-world. However findings in AR have been found to be highly variable, with no difference between AR and real-world having also been reported. The chapter consists of three studies examining the nature of AR depth perception. The first examines stereoscopic AR depth perception, followed by a test of relative cues (size and brightness) of virtual image in AR environment. The final experiment examines if depth information is derived from the virtual image or the real-world background. A two alternative forced-choice psychophysical perceptual matching task was implemented and designed to measure AR depth localisation. Participants' task was to judge if a virtual image (a green diamond) appeared either in front or behind a real-world target (pedestrian cut out). Depth thresholds for the AR image were tested with pedestrian target at 5m, 10m, 20m, and 25m at six locations relative to the pedestrian. The results from the first experiment revealed an effect of underestimation with depth error being less than previously reported in AR studies. The second experiment demonstrated that relative cues to depth for virtual imagery combine differently to real-world objects. The final experiment revealed that observers either acquire depth information from the virtual image or the real-world background, however not from both sources. Findings also highlighted the large individual differences and slow response times (above 2.5s average).

The current chapter explores depth perception in AR environments and is split into three sections; first assessing the ability of stereoscopic depth perception in an AR environment, secondly how virtual image depth cues (specifically relative brightness and relative size) combine and interact in an AR environment. Finally depth cues in the real-world scene are manipulated (linear perspective) on the impact they have on depth perception in AR.

Section A will begin by presenting previous experimental work assessing stereoscopic depth perception in virtual and augmented environments and real-world setting at distances in action space (1.5 – 30m). It will detail findings relating to virtual image location in terms of vertical and horizontal position and virtual images overlaid/occluding a real-world pedestrian. Followed by a novel methodology for measuring depth perception in AR will be presented. Results and analysis of the findings of stereoscopic AR depth perception will then be presented and discussed

Section B will present previous research in real-world environments assessing the impact of size and brightness in depth perception. This will be followed by a detailing of work in AR examining size and brightness and other cues to depth perception in AR. A set of three experiments will be presented, the first assessing brightness only cue, the second size only cue and finally brightness and size combined. Analysis of the results will be presented and discussed in terms of cue combination models.

Section C will investigate if depth information in an AR environment is acquired from the virtual image or the real-world background. The section will first discuss previous experimental work in AR environments assessing this question. An experiment, with the similar experimental design as in the previous experiments is presented, however the real-world background environment is manipulated (increased linear perceptive.

Section A: Stereopsis AR depth perception

3.1 Introduction

Perceptual equality in AR (between virtual image and real-world target) is important to achieve. For example AR devices such as vehicle HUDs, fast detection and correct responses to potential hazards could prevent serious accidents and damage to vehicles and drivers. A mismatch in depth between virtual image and highlighted object could cause increased errors in judgment and increased response times (Poitschke et al, 2008). An AR system which can accurately display depth cues by providing distinct focus distances for each visual element in the environment has been proposed as a genuine possibility for future AR devices (Gabbard, Fitch & Kim, 2014).

As discussed in '*1.11'* AR distance judgments have been measured in a variety of ways. A number of devices and methodologies have been implemented establishing different points of perceptual equality for virtual targets in real-world backgrounds.

The goal of 'Section A' is to present a new technique for measuring AR depth perception without the use of a see-through head mounted display (restricted field-of-view) and to establish observer's stereoscopic depth localisation thresholds at 5m, 10m, 20m and 25m. Findings will be compared to previous accounts of AR, VR and real-world depth estimates.

3.1.1 Comparative depth judgements in AR, VR and real-world settings

When comparing results between different environments it is important to consider the methodology used. Depth experiments in real-world settings, at medium to far distances, have implemented perceptual matching tasks, such as Allison et al (2009) and Palmisano et al, (2010) – see '1.11.1' for study details.

However in AR and VR studies this has not often been the case, as examinations of depth have often implemented HMD, in which using a perceptual matching task can potentially confound the results (Renner et al, 2013). Most studies in AR have used action-based methodologies and can produce different estimates.

Stereopsis depth perception in real-world settings for distance judgements between 1.5 – 30m (action space) has reported depth error at 0.5% of actual distance at 9m to 1% of distance at 18m (Allison et al, 2009). In studies examining distance judgments in virtual reality distance judgments, have found depth error ranges dramatically.

For example, Armbrüster, Wolter, Kuhlen, Spijkers & Fimm, (2008) examined depth in virtual environments at distances between 40cm – 500cm. A virtual target was presented in one of ten virtual distances between 40cm and 500cm; participants' task was to verbally estimate the distance. The depth error reported was at 21% of actual distance at only 5m (et al 2008). However this study did implement a verbal estimation task.

However examinations of stereopsis depth in AR have established an average of 15% error between distances of 11m – 27m (Swan et al 2007). Swan II et al (2007) experiment is one of few which implemented a perceptual matching task. As in conjunction with a HMD a perceptual matching task potentially confounds results. This is because HMDs completely block out reality, and when combining with a perceptual matching task only comparisons to remembered or virtual reference points are being made (Renner, et al 2013). To accurately compare real-world depth perception, at medium to far distances,

AR depth perception needs to be measured in as similar fashion as possible to real-world depth estimates.

The current study presents (see *3.2 Method*) a novel way of presenting virtual imagery without the use of a HMD or a built in see through display. Using reflection, the experiment design allowed for presenting virtual images accurately at a precise distance without a HMD or built in see through display to give an accurate depiction of stereoscopic depth perception in an AR environment. Whilst allowing for a perceptual matching task design similar to Allison et al (2009) and Palmisano et al (2010).

3.1.2 Vertical image location in the visual field

In an AR display virtual imagery may be presented at different heights in the visual field relative to the real-world object it is highlighting.

Height in the visual field is considered to be an effective depth cue in real-world environments. (Cutting & Vishton, 1995); research suggests this may be the same for AR environments (Kirkley, 2003; Swan II & Livingston, 2006). Gibson (1950) proposed that the horizon acts a reference point for perception of distance to objects in the environment. That is, objects that appear low relative to the horizon tend to be closer than objects that appear higher up in the visual field relative to the horizon. Further, Gibson (1950) highlighted the importance of the ground in depth perception. This is potentially because most objects in natural scenes are located below eye level (Yang & Purves, 2003); however only in action space.

Investigating human factors issues in AR environments Kirkley (2003) tested participants' distance perception for virtual and real objects in a real world environment between distances of 3 – 33.5m using HMD and a verbal report technique. Participant's performance significantly improved with objects

placed on the ground. An overestimation of depth was found for virtual imagery placed high in this visual field but on the ground plane participants underestimated depth. An inability of observers to scan the ground plane in a near-to-far field direction was given as a possible explanation for the observed depth underestimations.

Swan II & Livingston (2006) also investigated height in the visual field in their perceptual matching task for AR environments tested between 2.25m – 44.31m, reporting similar results to Kirkley (2003); an overestimation of distance was recorded when AR presented in the upper visual field and an underestimation of distance when presented in lower part of the visual field.

Investigations examining depth judgments for real and virtual objects at far distances (30m >) have found other interesting effects of height in the visual field. Wither & Hollerer (2005) conducted study to determine how well above real-world target depth cues performed in the absolute case at distances 20m – 80m. Results found height was less useful further away. Livingston et al (2003) explored the impact of relative height in a far-field occluded AR study. It was found that error increased when imagery was placed on the ground compared to when placed just off the ground. However relative height is not the same as absolute height, a line of research AR depth perception has not explored.

An important consideration to be taken into account for findings using AR HMDs is the limited FOV ($< 50^{o}$) in vertical directions; limiting ground perception. Wu et al (2000) discovered a loss in visibility for the ground at near distances due to the limited FOV, potentially affecting depth judgements; a finding of underestimation when the head was kept still. Knapp and Loomis (2003) did not find an effect of underestimation when participants, using HMD, were able to move their heads (near to far/ down to up). It was concluded the importance of directing gaze to the ground in AR environments; especially in

application with limited FOV. However these conflicting results may not have been about limited FOV, rather the effect of motion parallax.

In the current experiment virtual imagery was presented through a large see-through display, thus eliminating the issue of a limited FOV present in HMDs. Also a chin rest was used to eliminate any potential head movements from participants, therefore eliminating the additional depth cue of motion parallax.

3.1.3 Horizontal image location in the visual field

It has been shown that the horizontal position of a target affects stereoscopic depth perception in real world environments. Foley (1980) examined participant's depth perception (<1.8m) in a relative distance task; participants were required to align a target on their perceived apparent frontal plane. Results demonstrated that positional errors increased as a function of horizontal frontal distance from a centre point (Foley, 1980).

Head and eye movements have also been found to effect horizontal frontal position distance estimates. Arthur (2000) proposed guidelines for targets requiring being less than 10^{o} from a centre point, and only minimal to zero head movements involving small saccades. To limit head movements a chin rest was used in the current experiment. The examination of AR depth perception in the following experiment involved conditions where virtual imagery was both 5^{o} to the left and right of the centre real-world target. So far there has been no examination of relative horizontal positions for stereoscopic estimates of AR environments for distances upto 30m.

3.1.4 Overlaid virtual imagery

AR devices can highlight real-world targets by overlaying virtual imagery directly over the object of interest. For instance an outline of a shape can be

applied to a scene to direct a users' gaze. An outline of a shape rather than a full shape does not occlude the real-world target.

Experiments evaluating depth judgments in AR have often examined the effect of occlusion. Research has found that an occluding surface in a scene can increase the error of depth judgments between distances of 5 – 45m (Swan et al, 2006). However the inclusion of occlusion as a depth cue has found to improve depth judgments (Livingston et al, 2003; Kyto et al, 2013).

In the current examination of AR depth perception the virtual image will not be occluding the real-world target as the virtual target will be an outline of a diamond (see *3.2.2*). Testing will compare if AR depth estimates are improved for imagery presented adjacent to the real-world target or overlaid directly.

3.1.5 Summary and Rationale

Results and findings from stereoscopic AR depth perception have been inconsistent; for example the variance of depth error and distance underestimation been reported ranging from 6% - 68% of actual distance (Kirkley, 2003).

These inconsistencies may be due to the large use of methodologies; triangulated blind walking, verbal reports (Renner et al, 2013). As well as a large number of different AR devices used in AR depth perception; see-through display and HMDs (Grechkin et al, 2010). It is difficult to compare results in AR with those of real-world as different methods have been applied. Generally perceptual matching task applied to real world studies (e.g. Allison et al 2009; Palimsano et al, 2011). In AR and VR environments designs have used action based or verbal based reports (e.g. Jones et al 2011, Armbuster et al 2008).

To try and solve these issues a novel experimental method was designed to measure AR depth perception without the use of HMD or see-through

devices. The current study measured performance as AR images were presented at different distances from a real-world object. These depth thresholds were measured when the real object was at 5m 10m, 20m, and 25m. The intention was to develop clear indications to tell give an indication when an AR image is at a distance that is indistinguishable from the depth of a real object, examining relative height of an AR image with regards to real-world target.

3.1.6 Research Question & hypothesis

Is stereopsis depth perception in AR, when applying methodologies for real-world depth estimates (Allison et al, 2009), still produce similar error and underestimation as previously found in AR (Swan et al, 2007)?

- Hypothesis 1: Depth perception in AR will be poor in resolution compared to real-world
    - Prediction1.1: An underestimation of depth to be established, with judgments being significantly more than 1% error
        - Tested by comparing current findings of depth error

- Hypothesis 2: Depth perception will deteriorate as perceived distance increases
    - Prediction 2.1: Error in depth judgments will increase as a function of real-world target distance
        - Tested by main effects of ANOVA

- Hypothesis 3: Height in the visual field will have a significant impact on depth judgments
    - Prediction 3.1: depth estimates on the ground being significantly superior to presentation higher in the visual field.
        - Tested by main effects of ANOVA

3.2 Method

3.2.1 Participants

A total of 40 participants (16 male, 24 female) took part in the experiment. A mean age of 24.2 (SD = 3.65) with an age range of 20 – 37years. All participants were either students or members of staff at the University of Nottingham. A stereo acuity TNO test was administered to each participant, to make sure all had normal depth perception. The criterion for inclusion of the experiment was a score of ≤ 120 sec arc on the TNO; this indicated the participant did have the ability of stereopsis perception (Walraven & Janzen, 1993).

3.2.2 Experimental Design

The main goal of the study was to measure the ability of participants to judge the relative depth of AR and real-world objects. The independent variables examined were AR image position (horizontal and vertical position of the AR image in the visual field relative to the real-world object target, as well as the impact of occlusion cue).

Vertical position of an AR image at two locations; above the target and below, horizontal positions in the visual field of an AR image; left of the target, and right of the target. AR imagery was also overlapped on to the real world target.

Other independent variable measured was real-world target distance tested and at four representative distances; 5m, 10m, 20m and 25m. For the real-world target, a static 'pedestrian' was used; a cardboard cut-out of a man, which stood at 180cm in height (see *Figure 3.1*). The AR image used was a green diamond, which was presented at six different locations in depth relative to the static pedestrian target. Three locations of AR image were placed in front, and three

placed behind the static pedestrian target. Dependent variables were depth localisation accuracy and reaction times.

There were three presentations of the diamond at each virtual image distance, giving 18 presentations for each static target block. There were also three blocks of height in the visual field at each distance (top, middle, ground, or left, overlaid, right). A total of 216 AR diamonds were presented to participants over the course of the experiment. Response times and accuracy of depth judgment were measured.

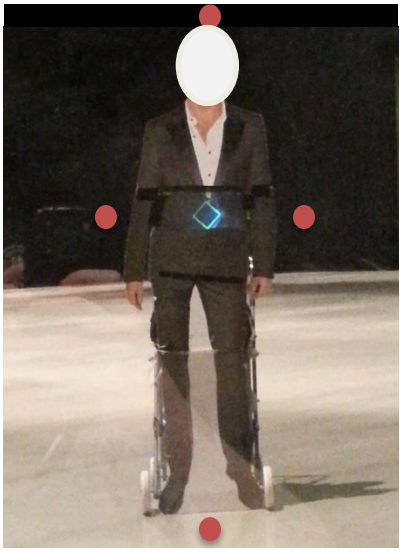


Figure 3.1: An outline of a green diamond was used as our AR image. The red dots represent the locations the AR image was presented relative to the static pedestrian target. The face has been hidden for publication and was not so in the experiment.

3.2.3 Showing the AR Image

To test AR depth perception in a vehicle-like configuration, the AR image was created by reflecting a projection of a green diamond carefully through a mock-up windscreen (see *Figure 3.2*). Using a projector and a silver screen, a reflection of a green diamond was directed into a mirror, and positioned to direct the reflection of the green diamond through the windscreen. The green

diamond appeared in an area in space either in front or behind the static pedestrian target. To manipulate the distance of the imagery - the projector and reflective screen were moved further away (or closer) to the mirror. The mirror was always 2.5m in distance away from the windscreen. So for example, for an image to be at 25m from the participant the silver screen, on which the diamond is projected on, would need to be 22.5m away from the mirror.

The augmented image was generated using Psychopy v1.81on ASUS X550C NVIDIA GeForce 720M laptop. A NEC PA550W 5500 LCD projector was used to project the green diamond.

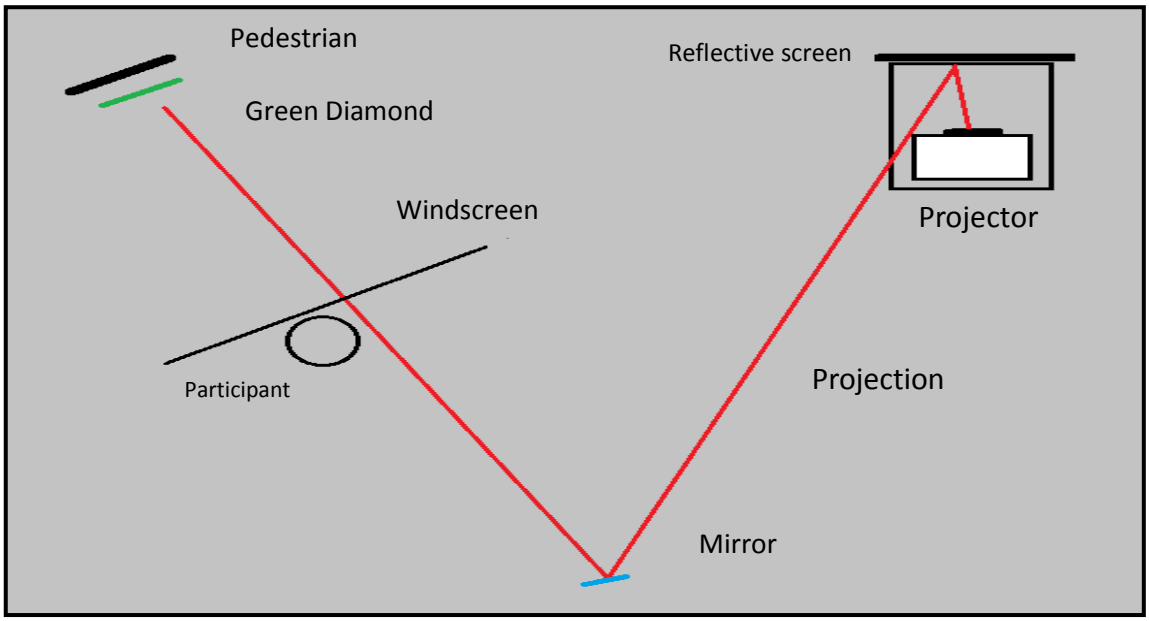


Figure 3.2: Schematic of experimental paradigm to produce AR image

3.2.5 Image height

The height of the image was changed by and increasing or decreasing the physical height of the mirror. A similar procedure was applied for horizontal image positions.

Due to the size of the mirrors we found that at very close distances it was not possible to manipulate the position of the green diamond, therefore at 5m, we only tested participants on the ‘overlaid’ position.

The experiment took place in a TV studio at the University of Nottingham, this allowed for a wide open space for testing (see Figure 3.4) with few additional depth cues in the environment. A black carpet was added to reduce shadowing cues on the floor.

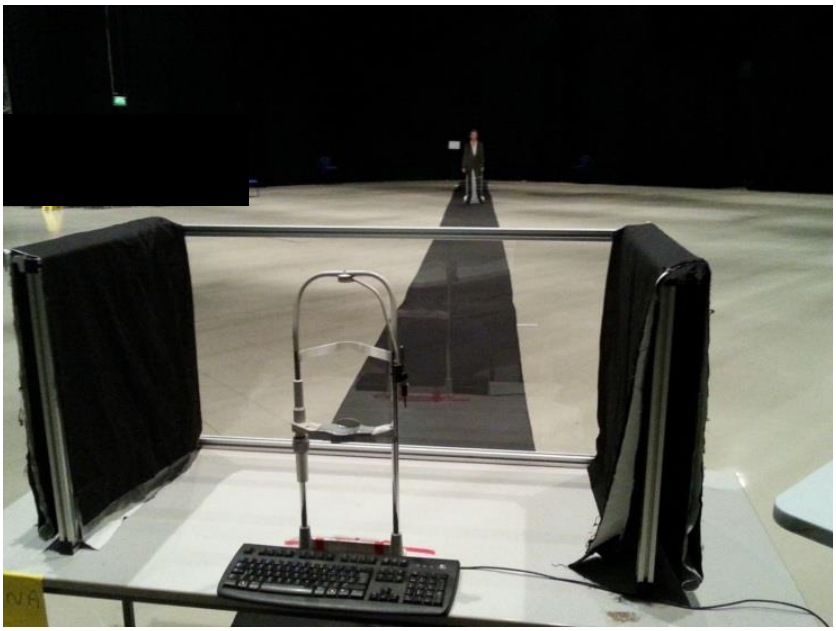


Figure 3.4: TV studio setting where AR depth perception experiment took place

3.2.6 Procedure

We expected that depth discrimination performance would change with distance. For this reason we scaled the range of AR image locations used for each participant and each distance tested based on initial pilot testing. Before the main measurements each participant completed an adaptive 3up 1down staircase at each distance of the real-world target. The locations of the AR image were then set to be 2.5, 1.0, and 0.5 multiple of the depth threshold recorded from the staircases.

For each trial the participants were asked if the green diamond was in front or behind the static pedestrian target. If they judged the diamond to be in front participants pressed ‘down’. If they judged the diamond to be behind they pressed ‘up’. Each trial began with a blank screen (or an empty field) for 10s, followed by a tone for 2s. This was followed by the appearance of the green diamond. The diamond disappeared once the observer had made their response, or once 10s had passed.

Participants were given two full practice trials to become accustomed to the imagery. For each position in depth, for the AR image, there were three repeats of the task. Each block took around 5 minutes. After each block the pedestrian was moved to a new distance. Different horizontal or vertical positions were presented in different (counterbalanced) blocks, with a short break for the participant in between. Each block of trials consisted of 18 diamonds. For the AR imagery positions, the distances for each block were completely randomized by the experimental software. On average each experiment lasted 90 minutes.

3.3 Results

Depth thresholds were obtained for every participant at each real-world target distance. Performance was measured in terms of accuracy and response times. Accuracy was plotted for each AR image location. A curve was fit to the accuracy data separately for each participant, at each distance and position. This curve was used to extrapolate the AR depth position where participants judge the AR image to be in front of the real-world object on 75% of the time, on 25% of the time (i.e. almost always judged behind) and where they are equally likely to report in front or behind (since there was no ’same’ option, this equivalent to being judged at the same depth). The point of perceptual equality was calculated for the AR image distance which was reported in front on 50% of

occasions. In figures *3.3, 3.4 and 3.5* example fitted curve functions for AR presentations matching the real-world pedestrian target distance are provided.

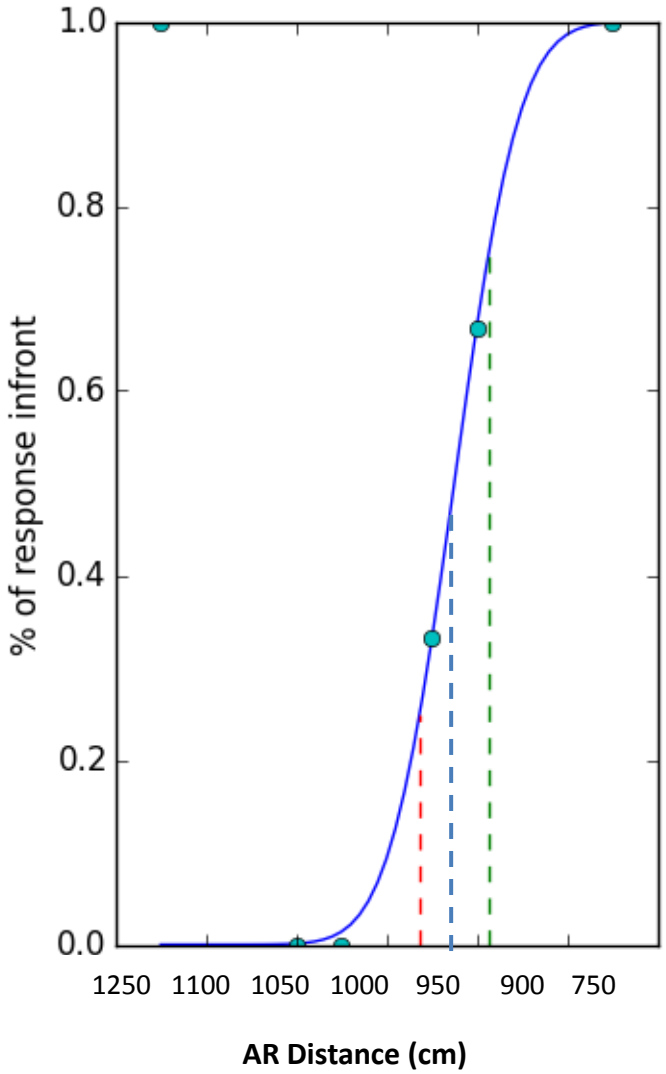


Figure 3.3: An example psychometric fitted function representing a participant who judged perceptual equality at 970cm (perceived infront) for virtual object matching a real-world target at 1000cm. On the Y-axis is the percentage of times a participant responded that the virtual target was perceived to be infront of the real-world target. On the X-axis is the distance of the AR image. Red dotted line represent 25% of the time participant responded infront, the green line 75% and blue line represent 50% (perceptual equality)

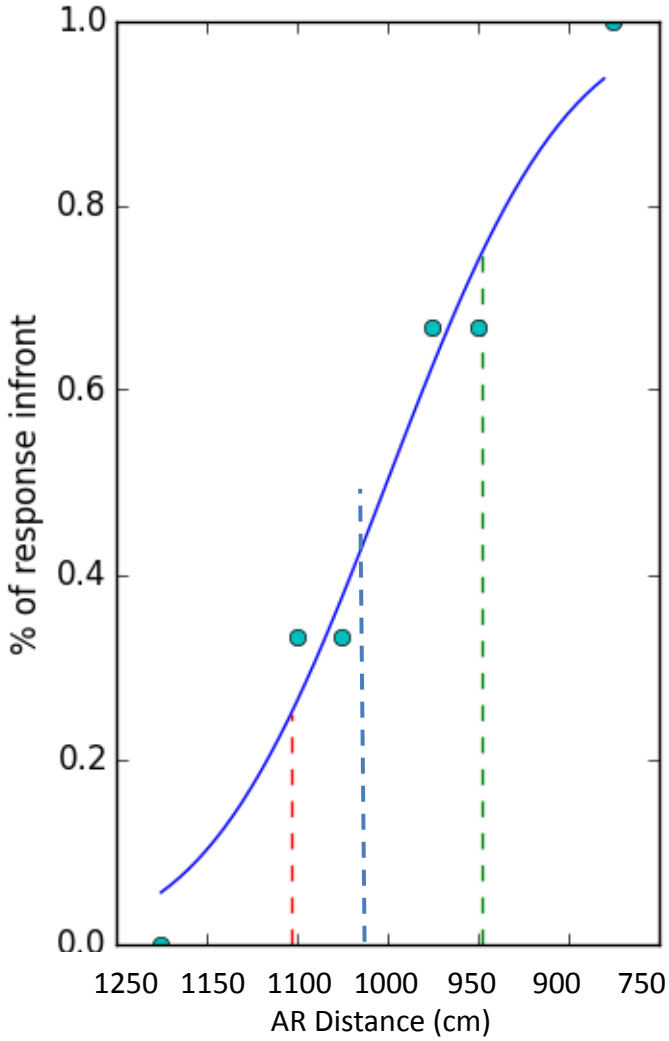


Figure 3.4: An example psychometric fitted function representing a participant who judged perceptual equality at 1000cm for virtual object matching a real-world target at 1000cm. On the Y-axis is the percentage of times a participant responded that the virtual target was perceived to be infront of the real-world target. On the X-axis is the distance of the AR image. Red dotted line represent 25% of the time participant responded infront, the green line 75% and blue line represent 50% (perceptual equality)

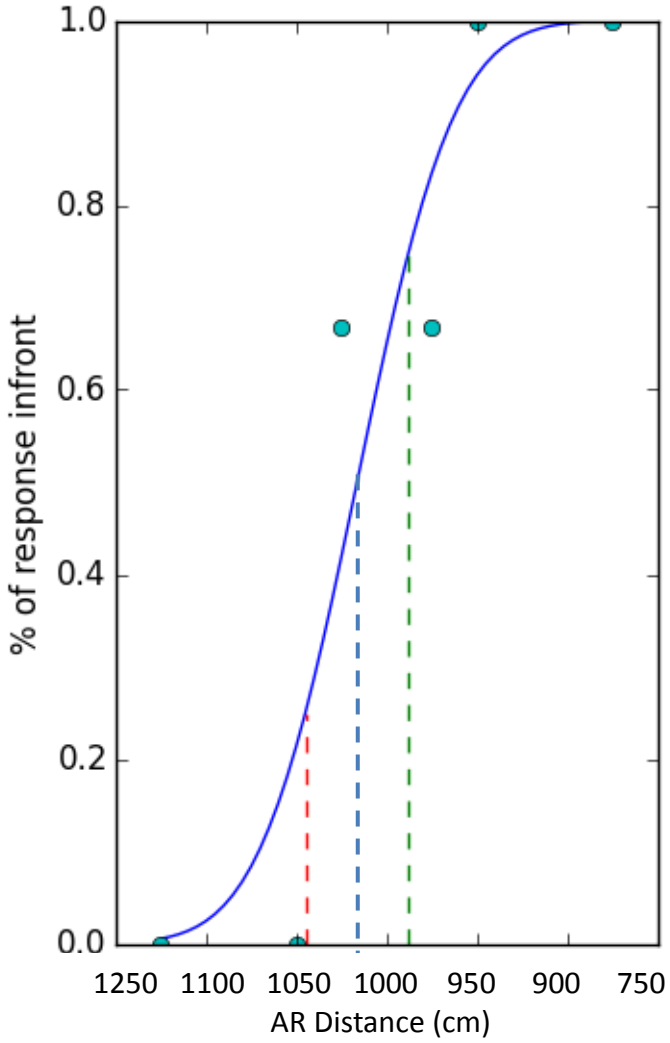


Figure 3.5: An example psychometric fitted function representing a participant who judged perceptual equality at 1020cm (perceived behind) for virtual object matching a real-world target at 1000cm. On the Y-axis is the percentage of times a participant responded that the virtual target was perceived to be infront of the real-world target. On the X-axis is the distance of the AR image. Red dotted line represent 25% of the time participant responded infront, the green line 75% and blue line represent 50% (perceptual equality)

3.3.1 Perceptual equality of AR depth perception

Analysis of participants psychometric functions at 50% of the time the virtual image was perceived to be in front of the real-world object (i.e. the distance where the virtual image was perceived to be at the same distance as the pedestrian), was analysed using a two-way 3 (Real-world target distance: 10, 20 and 25m) X 3 (Positon in visual field) within-subjects ANOVA. After assumptions of sphericity were reached, no interaction between height and distance was found, however a main effect of distance was established ($F(2,28)= 3.7$, $p = 0.015$, $\eta_p^2 = 241$). Bonferroni corrected pairwise comparisons revealed a significant difference between presentations at 10m and 25m ($p =0.025$), with presentations at 10m being superior to 25m.

For horizontal positons tested, another 3 x 3 within-subjects ANOVA was carried out. After assumptions of sphericity were reached, no interaction between height and distance was found, however main effects of both distance ($F(2,30) = 4.5$, $p = 0.019$, $\eta_p^2 = 232$) and position in the visual field ($F(2,30) = 5.1$, $p = 0.012$, $\eta_p^2 = .255$) were established. Post hoc analysis used Bonferroni corrected pairwise comparisons to reveal a significant difference between AR presentations overlaid on the real-world target compared to those presented to the right ($p = 0.024$). Projections overlaid on the real-world target producing more accurate estimate compared to presentations on the right of the real-world target. Bonferroni corrected pairwise comparisons also revealed a significant difference between 10m and 25m ($p = 0.015$), with presentations at 10m being superior to 25m.

The error of depth judgments was calculated (see Figure 3.4) at 6% (5m), 10% (10m), 9% (20m) and 11% (25m). This result is similar to the results of Swan II et al (2007) who found 15% depth error judgments for AR depth perception.

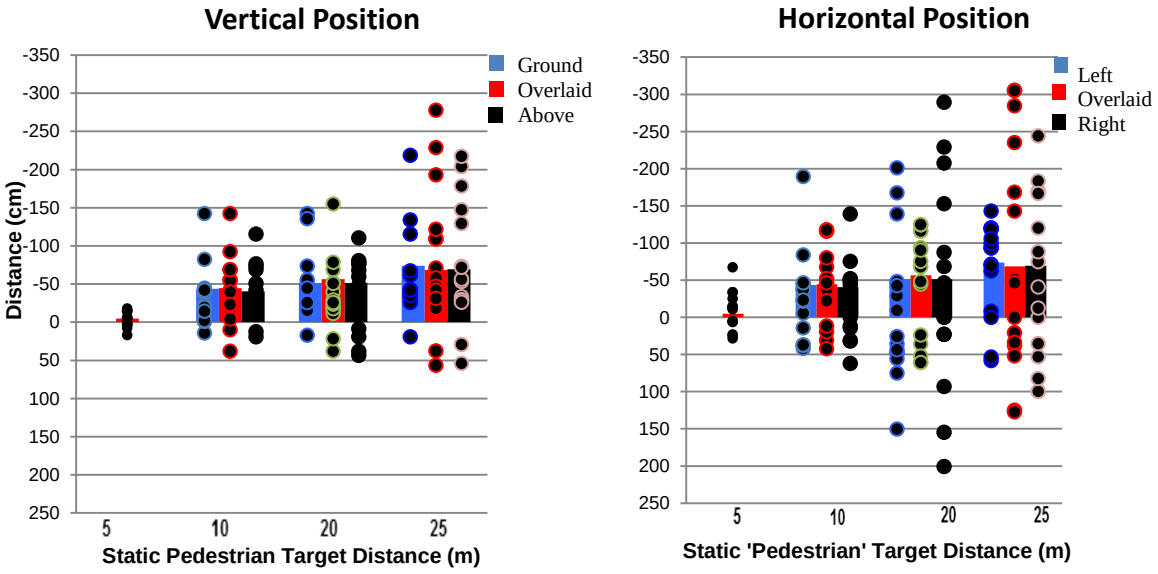


Figure 3.3: Left graph represents the point of perceptual equality of the AR image and real-world target for vertical image position experiment. The right graph represents the point of perceptual equality of the AR image and real-world target for horizontal image position experiment. The Y-axis is the AR image distance, with the X-axis representing the real-world target distances. The Bars represent the mean distance separation of AR and real-world target. Circles represent individual data scores**.**

3.3.2 Virtual image perceived in front

Analysis of the psychometric function at 75% of the time participants perceived the image to be in front (or rather the distance where the green diamond was perceived to be in front of the pedestrian), was analysed using a 3 (Distance) X 3 (Positon in Visual Field) within-subjects analysis of variance. After assumptions of sphericity were reached results established no significant interaction, however main effects of distance ($F(2,38)=9.8$, p <.0001 η_p^2 = .220), and of image height ($F(2,38)= 3.2$ p<.0001 η_p^2 = 115) were established. Post hoc analysis using Bonferroni corrected pairwise comparisons revealed significant differences between 10m and 25m (p= 0.03), with superior performance at 10m. Pairwise comparisons also revealed significant difference

at height with ground producing superior depth judgments than presentations overlaid (p =.009) and above the real world target (p = .002).

For horizontal virtual image positon, another 3 X 3 within-subjects ANOVA was carried out. Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(5) = .231$, p =.003, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .61). A significant interaction between distance and position was produced ($F(4,76) = 3.6$, $p = .027$, $\eta_p^2 = .160$). A main effect of distance ($F(2,38) = 11.5$, $p < .0001$, $\eta_p^2 = .377$) was found (see *Figure 3.5*). Bonferroni corrected pairwise comparisons revealed significant differences between 10m and 20m (p= .001), and 10m and 25m (p=.001) with superior performance at 10m (see *Figure 3.5*).

3.3.3 Virtual image perceived behind

Analysis of the function at 25% of the time participants perceived the image to be in front (or rather distance mostly likely to be seen behind the pedestrian), was analysed using a 3 (Distance) X 3 (Positon in Visual Field) within-subjects ANOVA. After assumptions of sphericity were achieved results found no significant interaction, however a main effect of distance was established ($F(2,38)=3.64$, p = 0.045 η_p^2 = .151). Bonferroni corrected pairwise comparisons revealed significant differences between 10m and 20m (p= .001), and 10m and 25m (p=.001) with superior performance at 10m.

Analysis for horizontal positon conditions of the AR image found no significant interaction of horizontal position and distance after assumptions of sphericity were achieved. A main effect of distance ($F(2,28)=9.6$, p = 0.001 η_p^2 = 321) was established (see *Figure 3.6*). Post hoc testing using Bonferroni corrected pairwise comparisons revealed significant differences between 10m

and 20m (p= .001), and 10m and 25m (p=.001) with superior performance at 10m.

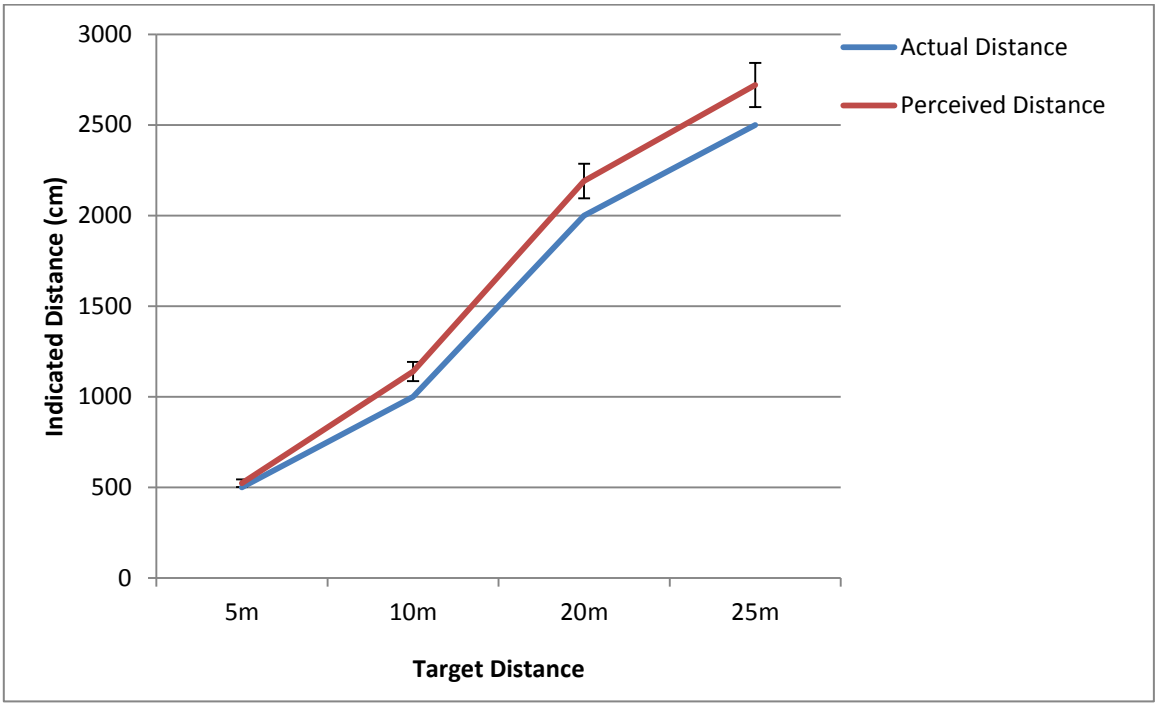


Figure 3.4: Graph representing the distance judged by the participants of the virtual image (red) compared to the actual distance of the real-world target. Error bars represent

3.3.4 Slope Analysis

Further analysis compared the slope of the psychometric function for participants. This measurement gives an indication of participant's sensitivity to change in the virtual image stimulus. A two-way 3 X (Pedestrian Distance) 3 (Positon in Visual Field) within-subjects ANOVA for vertical positons, after assumptions of sphericity were achieved revealed no significant interaction, however a main effect of distance was found ($F(2,28)= 18.6$, $p < .0001$, $\eta_p^2 = .495$). Bonferroni corrected pairwise analysis revealed significant differences between 10m and 20m ($p = .001$), and 10m and 25m ($p < .0001$) with superior performance at 10m.

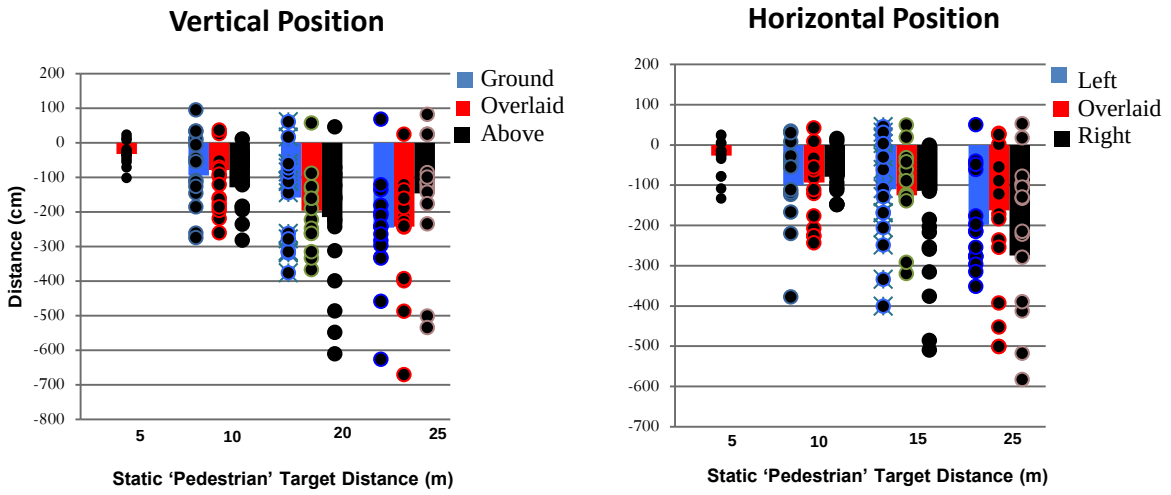


Figure 3.5: 75% proportion of responses where the AR image was 'in front' of the pedestrian. Left are results for the vertical presentations of image, right are the horizontal presentations of AR image. The circles represent individual data scores

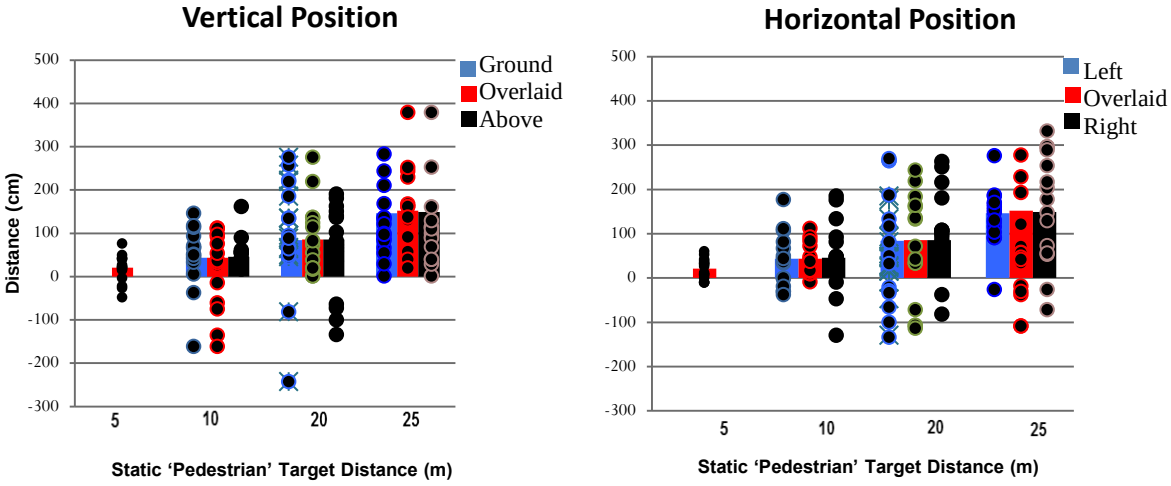


Figure 3.6: 25% proportion of responses where the AR image was 'in front' of the pedestrian. Left are results for the vertical presentations of image, right horizontal presentations of AR image. The circles represent individual data scores

For horizontal condition a 3 x 3 within subject ANOVA was carried out. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(9) = .934$, $p < .0001$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .53$). No significant interaction was discovered, although a main effect of horizontal image position was established ($F(2,28)= 22.7$, $p < .0001$, $\eta_p^2 = .545$). Bonferroni corrected pairwise analysis revealed significant differences between 10m and 20m ($p = .008$), and 10m and 25m ($p < .0001$) with superior performance at 10m (see *Figure 3.6*)

These results further maintain AR image needed to be closer to the pedestrian at distances nearer the observer for them to be indistinguishable in terms of depth.

3.3.5 Reaction Times

A two-way 3 (Pedestrian Distance) X 3 (Positon in Visual Field) within-subjects ANOVA for participant response times was carried out. After assumptions of sphericity were reached there was no interaction or main effects established (see *Figure 3.7*).

For horizontal positon a two-way 3 X 3 within-subject ANOVA did not find a significant interaction (after assumptions of sphericity were met). No main effects of distance or position were established. Response times were not significantly faster at close range real-world object distances (10m) compared to far-range (20 and 25m) real-world target distances (see *Figure 3.7*)

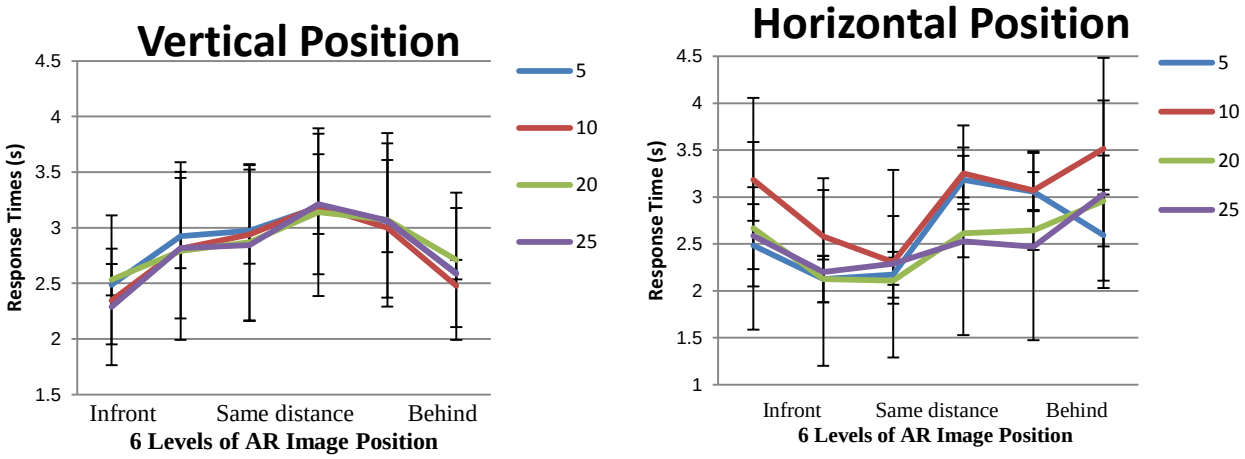


Figure 3.7: Average response times for participants depth judgments with AR image. Left on X-axis represents presentations in front of the pedestrian, and right represents presentations behind the pedestrian, with the points in between representing positions closest to real-world object. Error bars represent S.E.M

3.4 Discussion

The first study of the chapter presents two experiments investigating stereo acuity depth localisation for a virtual image in a real-world environment. Previous research in depth perception for AR environments have used a variety of measuring methods and devices to produce virtual images; with a variety of findings and conclusion with regards to underestimation and error (Kirkley, 2003). A novel design for measuring AR depth perception was carried out implementing a perceptual matching task for real world target distances of 5, 10, 20 and 25m. Results aimed to clarify the point at which virtual image is perceptually at the same distance as a real-world target for binocular disparity, and to compare with real-world depth estimates.

3.4.1 Depth error in AR

Analysing the depth error at each distance and subsequently comparing with previous results from AR studies, it was revealed that a similar error in

depth judgments as previous research was found. In the current study depth error (in action space) averaged at 13% of the actual distance of virtual image.

In comparison to similar studies using a perceptual matching task measuring stereoscopic depth perception in virtual environments, depth error was recorded at 15% error (Swan II et al, 2007). When comparing the current results to real-world depth estimates using binocular disparity, depth perception appears to be inferior, with reports of only 1% error at upto 18m (Allison et al, 2009).

Results in the present study suggest depth error, when measured comparability using perceptual matching task to be around 10 – 15% of its actual distance. The results are similar to those found previously in AR using perceptual matching (e.g. Swan et al II, 2007). AR depth perception seems to be superior to VR and inferior to real world. This may indicate that in AR depth perception the virtual reality elements are behaving such as virtual reality, i.e. unreliable and unstable depth information being derived. The real-world information derived is stable and reliable. This is why depth judgments in AR environments appears to in between performance of VR and real world.

However in the current study key depth cues of relative size and brightness of the virtual image were not available for participants. It will be interesting to find out when these cues are available for participants (in the follow up experiment of the chapter) if depth perception in the AR environment matches that of real-world estimates.

3.4.2 Underestimation in AR

Past research on depth perception in AR has produced inconsistent results with regards to an underestimation or overestimation of depth (Kirkley, 2003; Jones et al, 2008; Swan et al 2007). Examining the distances where observers could not discriminate between the depths of the AR image and the real-world

there was a small amount of underestimation ranging from 3% at closest distance of 5m, and 12.5% at 25m. This supports some of the previous findings of only a small difference in AR depth judgments compared to real-world depth (Jones et al, 2008).

Results revealed that depth judgements increased with increasing distance from the participant; the farther away the real-world target, the larger the separation between image and object can be for a participant to see them both at the same location in the distance. The further away the target the larger the error. These findings replicate both real-world and virtual environments (Mckee & Taylor, 2010, Swan et al, 2007; Jones et al, 2008).

3.4.3 Virtual image position

An effect of the position of AR image in the visual field (height and horizontally) was found when participants distinguished an image to be infront of the pedestrian target. This was not the case for depth equality results; suggesting AR imagery could be placed at the same depth, in relation to real-world objects, for all locations in the visual field. Imagery presented on the ground produced the same error as imagery presented above the real world target. This is a contradictory finding to previous research (Kirkley 2003; Swan et al, 2007). For instance (Kirkley, 2003) established a main effect of height in AR judgments, with (Jones et al 2007) revealing depth performance increased with imagery on the ground.

For the experiment presented, in one of the positional conditions virtual imagery was overlaid on the real-world target. The virtual image used was not a solid shape, an outline of a diamond, thus did not occlude the real-world target.

Comparing the overlaid position to presentations of non-overlaid image positions no significant differences in depth equality was established; a result

which could have a risen as participants were only able to use stereo acuity rather than the size and brightness of imagery as a cue. The virtual image used in the current study was only an outline of a diamond. Past research has used solid shapes, which may have affected how AR is perceived in depth when overlaid onto a real-world object compared to a position adjacent to an object.

3.4.4 Individual difference & slow response times

Large individual differences were established and slow response times (above 2.5s average), illustrate the possible difficulty involved when judging AR in depth. This is a concern when applying AR to technological devices. For instance if drivers spend a significant amount of time interpreting the depth of AR HUD imagery, it could potential demand too much of a driver's attention. It is possible that then these slow response times were due to our controlling for cues, such as relative size and brightness, in the study. It will be interesting to see how, or if, performance changes with added depth cues. The individual points on figures 4a and 4b illustrate the large individual differences in AR depth perception. This is an issue that HUD designers may need to take into account when implementing AR in vehicle HUDs. This may mean considering the possibility of having AR HUDs individually calibrated for every driver in terms of depth presentation of imagery.

3.4.5 Summary and Conclusion

Combining threshold data from two experiments this paper has presented findings for AR depth judgments, relative to a real-world target at 5m, 10m, 20m and 25m. Results have been presented in terms of actual distance where an image will been seen in front of a target, behind the target and at the point in depth where a user will perceive an AR image and a real-world object to be at the same distance (perceptual equality).

Findings revealed depth thresholds scaled with distance (larger thresholds at further distances, as well as highlighting large individual differences and slow response times (above 2.5s average). The result demonstrated depth localisation in AR to be similar to previous experiments for AR depth perception implementing perceptual matching task (Swan et al, 2007) compared to real world depth measurements (Allison et al 2009).

In conclusion, AR images appear to be poorly resolved in depth. Follow up studies will investigate the effects of additional of cues to depth for AR imagery depth localisation, such as relative size and brightness, and a combination of cues.

Section B: Relative cues of size and brightness in augmented reality depth perception

3.5 Introduction

Interpreting any distant features correctly in the real-world requires constructing an accurate depth estimate based on a limited number of natural cues. In AR environments the number of available natural cues is even fewer, although superior to that of VR; both being described as cue poor environments (Landy et al 1995). The current section of the chapter investigates virtual image cues of relative brightness, relative size and the effect on depth perception when combined. Testing will establish the overall effectiveness of each cue and a combination of size and brightness cues to AR depth localisation and examine how this may differ with real-world depth perception.

3.5.1 The importance of size and brightness depth cues in virtual and augmented environments

An important question for an AR system is which cues work when the user is being shown virtual representations of objects integrated into a real scene (Livingston et al, 2003).

Psychophysical testing has examined relative cues to depth for both size and brightness in 3D virtual environments. For example Surdick et al (1994) compared depth cues of relative size and brightness (amongst other depth cues; relative height, linear perspective, foreshortening, texture and stereopsis) using a method of constant stimuli for distances between 1 – 2m in a virtual environment. A flat 5cm square was presented floating in a virtual scene. The participants' task was to indicate which one of two stimuli was presented closer.

Results revealed relative brightness to be the least effective cue to depth (a small change in specific depth cue results in a noticeable change in perceived depth (Surdick et al, 1994)), and significantly inferior to all other cues at 1m. However this was not found to be significantly inferior to relative size at 2m (Sedgwick et al, 1994). Relative size and brightness, as well as stereopsis, were the only cues tested which were significantly more effective at 1m than at 2m. The researchers suggested that as the relative brightness of an object drops off with the square of the distance from the point of light source, a large change in distance will only produce a small change in brightness (Sedgwick et al, 1994).

In a follow on study Surdick et al (1997) examined the ability to accurately present distance in virtual and augmented reality systems. Separately the effectiveness of a range of visual depth cues was tested; relative brightness, relative size, relative height, linear perspective, foreshadowing, texture gradient and stereopsis, as well as combining them all in one condition. Experimental testing distances were short, with testing not going beyond 2m.

Participants where shown greyscale pixel images viewed dichoptically (e.g. 'stimulus A' presented in left eye and 'stimulus B' presented in the right eye) of two simulated fronto-parallel flat squared surfaces that appeared to float in space. The scene was a wall and a floor where cues for depth were manipulated or removed. Participants had to decide which stimulus appeared closer.

This is a method which was adopted in the series of experiments conducted for this chapter; a background scene with minimal cue information, and a stimulus setup where the cues to depth could be removed or manipulated. However the present research was conducted over much larger distances (4m – 30m of virtual image distance), with the virtual image used being of higher

resolution and brightness; which could have a huge impact for how depth cues are interpreted and combined in AR.

Surdick et al (1997) found that the effectiveness of size and brightness both decreased as a function of viewing distance. Again the authors proposed this was because relative brightness decreases with the square of the object's distance from the light source. The decreased effectiveness of relative size was explained by the fact a smaller change in retinal image size of the square of an object occurs when a given change in a simulated distance of a stimulus at 2m compared to 1m viewing distance.

Results also revealed the that a condition which combined all the depth cues together was produced superior depth estimates compared to all other conditions tested. However this result was not significantly superior to the condition with perspective cues which exhibited JNDs similar to the all cues condition (Surdick et al 1997). A ceiling performance effect was given as a possible explanation for the results, as the best level of performance had already been obtained with just the perspective cues. Alternatively participants may have chosen only one (or a few) of the available cues to attend to; potentially indicating 'a winner takes all' model for combining cues.

Larger JNDs would have been expected to be found when depth cues were tested individually. Verbal reports from participants also suggested only one or a few cues were attended to, with the authors also noting that stereopsis may have played a key part (Surdick et al, 1997). Researchers concluded that combing all-cues to depth does not necessarily enhance distance perception.

Although Surdick et al (1997) presented some interesting, and useful findings, the experimental setup was very rudimentary. Although allowing for easy manipulation and control of depth cues, the experimental stimuli consisted of simple greyscale pixelate imagery on augmented on flat squared surfaces.

Augmented reality devices have a lot more information now, with much more complex imagery and background scenes full of rich depth information. Testing only took place between 1 – 2m, therefore it is difficult to extrapolate findings to work with AR perception at long distances. As we know from Cutting (1995), cues have different levels of impact over distance, therefore testing how cues may differ in quality over distance in an environment is still to be fully explored.

Wither & Holler (2005) conducted a study to determine how well depth cues above a real-world target in the absolute case (i.e. when there are no other annotations in the user's field of view) and relative case. In an outdoor experimental setting, real-world targets were placed between distances of 20 – 80m. Participants were required to place a spherical marker at the correct location of a real-world target.

The study compared different cues to aid the user in placing annotations at the correct distance for user; size, shadows, top down view and surrounding targets with changing colours. The authors discovered relative cues were a great assistance in determining depth, with error values approximately cut by half. Surprisingly to the researchers, shadow planes and top down views were not a significant help to users in refining their distance estimates significantly compared to size-only cue. The cause of this result was attributed to visibility limitations of the see-through screen used. An interesting result was the size only cue produced the fastest reaction times.

Sands and Lawson (2003) investigated the influence of additional depth cues may have on an AR system. In an AR cursor positioning task, AR users had to align a 3D virtual cursor with real-world object (wooden dowels) in a stereoscopic display. A number of depth cue conditions were tested such as

scale of known size (the scale comprised a series of black and white bands each 1cm apart).

Results indicated that all of the depth cues tested improved AR accuracy, but suggested that certain depth cues are more helpful than others, for example correctly orientated shadowing on imagery had the greatest improvement for conditions tested; a theory which coincides with maximum likelihood estimation.

It is important to consider previous work testing in AR were users have full control of the device. For instance beneficial information is presented to users (e.g. knowledge about positional height, and other prior information) and results could only be useful for the experimental conditions chosen. Threshold criterion will have possibly differed between participants, with judgment criterion changing across trials, causing the measured threshold to considerably vary. Also anticipation effects (when participants set stimuli levels based on previous trials rather than on their current sensory impression) need to be taken in to account when assessing results of AR depth perception exploring cue combinations, such as in Sands & Lawson (2003), Surdick et al (1997).

Previous research examining depth perception using optical see-through displays and perceptual matching task in augmented environments have not manipulated or isolated relative cues to depth perception at distances upto 25m.

3.5.2 Summary & Rationale

Examinations of depth perception for real world environments have explored how cues to depth combine together. Investigations in to AR have not fully explored the impact cues and combinations of cues have in an AR environment at intermediate distances.

The key difference between augmented and real–world is the presence of the virtual imagery in the foreground, with the background possessing the same natural cues as the real-world. Therefore the following section presents three experiments measuring the performance of AR depth localisation in three different virtual image conditions; size only, brightness only, and a combination of size & brightness.

3.5.3 Question & Hypothesis

Do relative depth cues of in AR interact combine in the same was as real-world depth perception?

- Hypothesis 1:
  - Prediction: Relative size will be more effective than relative brightness as a cue for depth
    - Tested by: ANOVA comparing two conditions, post hoc test revealing significant differences between conditions.

- Hypothesis 2: Combined cues to depth
  - Prediction: A combination of both size and brightness will produce the best depth performance, compared to size and brightness presented singularly
    - Tested by: ANOVA comparing two conditions, post hoc test revealing significant differences between conditions.

The following method and results section reports three experiments assessing size, brightness and a combination of size and brightness as cues to depth in AR.

3.6 Method

3.6.1 Participants

3.6.1.1 Size

A total of 21 participants (8 male, 13 female) took part in the size cue experiment. A mean age of 22.7 (SD = 2.15) with an age range of 18 – 27 years

3.6.1.2 Brightness

A total of 21 participants (9 male, 12 female) took part in the brightness cue experiment. A mean age of 23.1 (SD = 2.2) with an age range of 18 – 29years

3.6.1.3 Size and brightness combined

A total of 21 participants (7 male, 14 female) took part in the size + brightness cue experiment. A mean age of 23.7 (SD = 2.1) with an age range of 19 – 29years.

All participants in each experiment were either students or members of staff at the University of Nottingham. A stereo acuity TNO test was administered to each participant, to make sure all had normal depth perception.

3.6.2 Design

The three experiments all implemented the same within-subjects design as the experiments presented in Section A (see 3.2.3).Within each depth cue condition group, the other IVs tested were the height of the virtual image (Above and Overlaid) and real-world target distance (5, 10, 20 and 25m)**.** Dependent variables were depth localisation accuracy and reaction times.

3.6.3 Virtual image manipulations

In the size experiment, brightness was controlled for with participants being able to judge the distance of the virtual image using the cue of size. The brightness of the image needed to remain the same at every distance the virtual image was placed. For this brightness measurements of the projector were carried out at 5m, 10m, 20m and 25m. A linear trend was applied and reversed to brightness of the image therefore brightness always appeared to be the same and not dependent on the distance.

In the 'brightness' condition, size was controlled for with participants being able to judge the distance of the virtual image using changes in brightness. The size of the AR image was required to always appear the same size at different distances. For this the following equation was applied:

$$atan^2/\frac{s}{d}$$

Equation 3.1: Equation for presenting the virtual image to appear to be the same size at each distance judged

In the equation atan^2 represents the inverse tangent, where *s* represents size of the virtual image and *d* is the distance of the image. In the final experiment participants had both size and luminance cues of the virtual image.

3.6.4 Procedure

Participants were given two full practice trials to become accustomed to the virtual imagery. For each position in depth, for the AR image, there were three repeats of the task. Each block took around 5 minutes. After each block the pedestrian was moved to a new distance (10m, 20m, 25m). Different virtual image positions were presented in different (counterbalanced) blocks, with a short break for the participant in between. Each block of trials consisted of 18

diamonds. For the AR imagery positions, the distances for each block were completely randomized by the experimental software. On average each experiment lasted 90 minutes. The same was procedure was carried out for size cue, brightness cue and size and brightness cue.

3.7 Results

3.7.1 Which depth cue produced the best depth estimates?

Depth thresholds were obtained for every participant at each real-world target distance. The point of perceptual equality was calculated by the AR image distance which was reported in front on 50% of occasions which will give the perceptual equality distance of the AR image and real world target.

To establish the possible impact different depth cues have on AR perception a three-way 3 (Depth Cue: size, brightness, size and brightness) x 3 (Height: overlaid, above) x 3 (Distance: 10m, 20m, and 25m) between subjects ANOVA was carried out. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(2) = .856$, $p = .012$, therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity ($\varepsilon = .87$). A significant three-way interaction was established ($F(4,116) = 6.4$, $p < .0001$, $\eta_p^2 = .180$), with a main effect of depth cue being found ($F(2,58) = 4.9$, $p = .011$, $\eta_p^2 = .144$). No main effects of height or distance were established. This may have occurred because % of distance rather than absolute distance was examined.

Bonferroni corrected post hoc analysis reported significant differences between depth cues of brightness compared to and size + brightness cues ($p = .016$), with size + brightness producing superior depth thresholds. When comparing depth cues tested in this study, with the baseline depth threshold reported in the previous section of this chapter, only a combination of size and brightness produced significantly superior depth thresholds compared to the

stereo only baseline (p = .002). Brightness cue was not significantly different to size cue, but did produce the highest depth error overall.

To further explore the three way interaction, analysis was split into each experiment assessing a cue or combination of depth cues.

3.7.2 Experiment 1: Size

From the first set of analysis it was found that significant differences between cue conditions occur in AR. However do depth cues interact with distance and image height in the same way. To find out first of all a two-way 3 (height: ground, overlaid, above) x 3 (distance 10m, 20m and 25m) within subjects ANOVA was carried on the size only depth cue (see Figure 3.7). A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(9)$ = .157, p <.0001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .56). No significant interaction was established, however main effects of distance (F= (2, 38) = 4.4, p < .020, η_p^2 = .178) and height was established (F= (2, 38) = 6.9, p < .003 η_p^2 = .210). Main effects indicate depth judgements are increasingly more erroneous the further away from the observer (10m = 32cm; 20m = 63cm; 25m = 81cm all behind of the real world target). Bonferroni corrected pairwise comparisons revealed significant differences between 10m and 25m real-world target distances (p = .004). The smallest average estimates are produced when imagery is overlaid (26cm in front), significantly different compared to ground at 73cm behind (p = .002) and above, 76cm behind (p = .002).

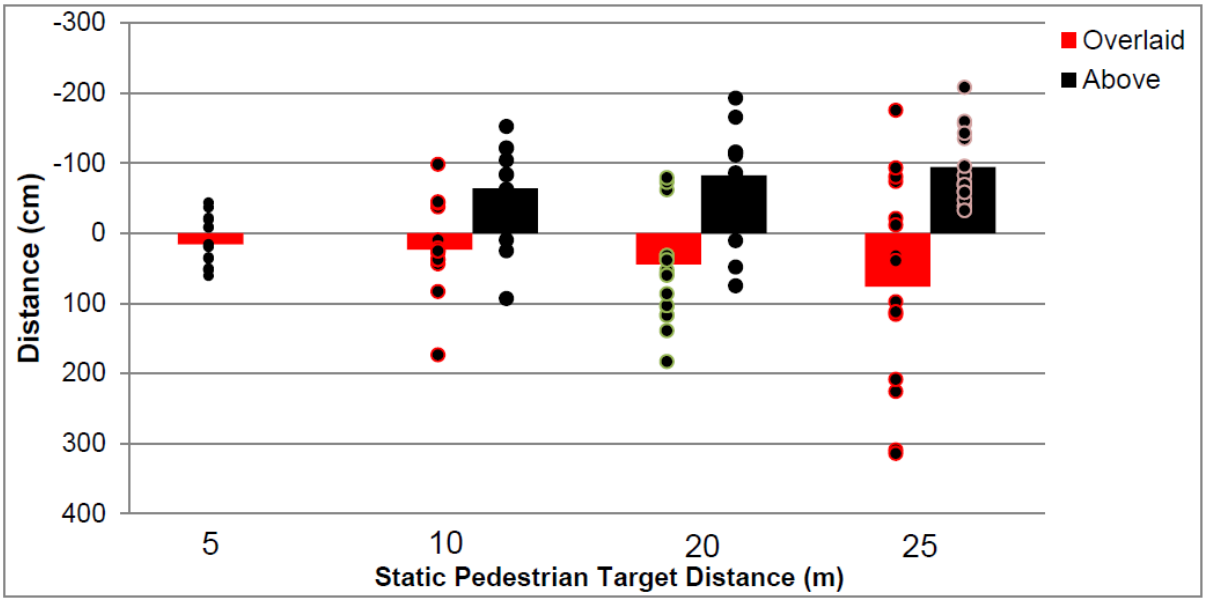


Figure 3.7: The point of perceptual equality of the AR image and real-world target for size only depth cue. Bars represent the mean distance separation of AR and real-world target. Circles represent individual data scores

3.7.3 Experiment 2: Brightness

To discover how the brightness of an AR image performs over distance and height, a two-way 3 (height: ground, overlaid, above) x 3 (distance 10, 20 and 25m) within-subjects ANOVA was carried out to assess the accuracy of AR depth perception at perceptual equality (see *Figure 3.8*). After assumptions sphericity were reached no significant interaction was found, however main effects of distance (F= (2, 38) = 7.7, p < .002, η_p^2 = .217) and height were established (F= (2, 38) = 5.0, p < .012, η_p^2 = .109). Main effects indicate depth judgements are increasingly more erroneous the further away from the observer (10m = 40cm; 20m = 43cm; 25m = 88cm behind of the real world target), with Bonferroni corrected pairwise comparisons indicating significant differences between 10m and 25m (p = .01). The smallest average depth estimates are produced when imagery is above (35cm behind); significantly different to overlaid (80cm behind) position (p = .01). A different finding compared to size only virtual image cue where occlusion produced the smallest depth estimates.

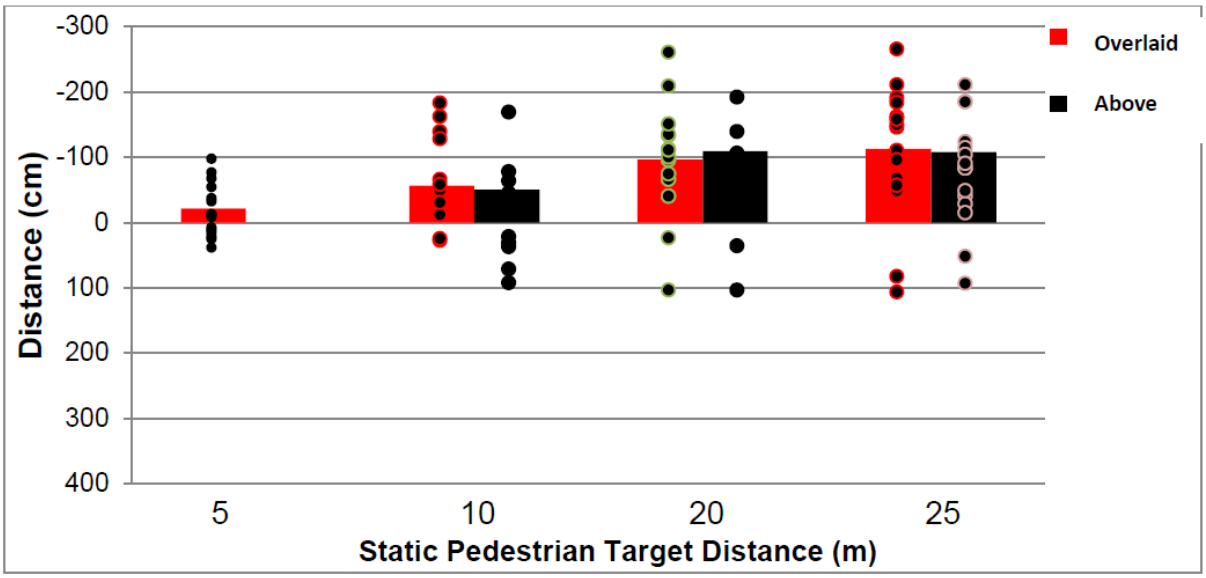


Figure 3.8: The point of perceptual equality of the AR image and real-world target for brightness only depth cue. Bars represent the mean distance separation of AR and real-world target. Circles represent individual data scores

3.7.4 Experiment 3: Size & Brightness

To find out if a depth perception of a virtual image is perceived differently when size and brightness cues are combined together, a two-way 3 x 3 within-subjects ANOVA was carried to assess the accuracy of AR depth perception at perceptual equality (see *Figure 3.9*). After assumptions sphericity was reached no significant interaction was established. However, a main effect of image positional height was established (F= (2, 38) = 21.6, p < .0001). The smallest average estimates are produced when the imagery is overlaid (29cm in font) compared to above (86cm behind) position. Bonferroni corrected pairwise comparisons indicating significant differences between 10m and 25m (p = .001).

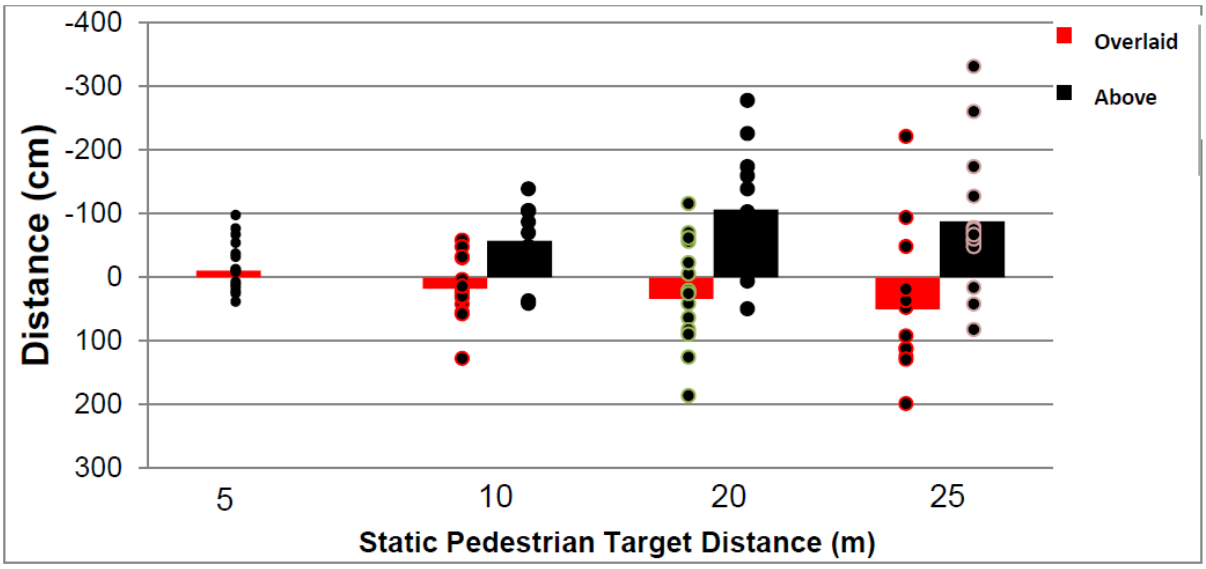


Figure 3.9: The point of perceptual equality of the AR image and real-world target for when combination of size and brightness depth cues. Bars represent the mean distance separation of AR and real-world target. Circles represent individual data scores.

3.7.5 Slope analysis

To establish if different cues to depth result in a greater sensitivity to change for distance estimates for the AR image, a three-way 3 (Depth Cue: size, brightness, size and brightness) x 2 (Height: overlaid, above) x 3 (Distance: 10m, 20m, and 25m) between subjects ANOVA was carried out. After assumptions sphericity were met, no significant three-way interaction was established, however a main effect of height was found ($F(2,114) = 35.5$, $p < .0001$, $\eta_p^2 = .384$). Additional slope analysis was divided into depth cue experiments; size, brightness, size + brightness.

3.7.5.1 *Size*

A two-way within 3 (Pedestrian Distance) X 3 (Positon in Visual Field) within-subjects ANOVA for size condition was completed using slope output. A Mauchly's test indicated that the assumption of sphericity had been violated:

$X^2(9)$ = .245, p =.004, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .63), no significant interaction was found, however a main effect of distance was found ($F(2,38)$= 11.8, p < .0001, η_p^2 = .381). Bonferroni corrected pairwise analysis revealed significant differences between 10m and 20m (p = .031), and 10m and 25m (p <.0001) with superior performance at 10m.

3.7.5.2 *Brightness*

A 3 X 3 within-subjects ANOVA for brightness condition was completed using slope output. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(9)$ = .290, p =.011, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .57), no significant interaction was found, however a main effect of distance was found ($F(2,38)$= 12.8, p < .0001, η_p^2 = .402). Bonferroni corrected pairwise analysis revealed significant differences between 10m and 20m (p < .001), and 10m and 25m (p <.0001) with superior performance at 10m.

3.7.5.3 *Size and Brightness*

A 3 X 3 within-subjects ANOVA for size and brightness condition was completed using slope output. After assumptions of sphericity were reached no significant interaction was found, however a main effect of distance was found ($F(2,38)$= 22.2, p < .0001, η_p^2 = .539). Bonferroni corrected pairwise analysis revealed significant differences between 10m and 20m (p < .001), and 10m and 25m (p <.0001) with superior performance at 10m.

3.7.6 Reaction times

Participant's response to depth judgements was recorded and initially analysed using a four way between subjects ANOVA was carried out examining the effect of the height of the virtual image 2 x (overlaid and above) of AR

imagery in relation to a real-world target, 3 X distance (10m, 20m, 25m) 3 X multiple of threshold (2.5, 1.0, 0.5) X 2 placement of AR image (infront or behind).

After assumptions of sphericity were achieved a significant four-way interaction was established ($F(4,228) = 11.7$, $p < 0.0001$, $\eta_p^2 = .170$), with a main effects AR position in front or behind the real world target ($F(1,57) = 17.6$, $p < .0001$, $\eta_p^2 = .236$), with AR positioned in front (2.4s) of the real-world target being significantly faster than when projected behind (2.7s). Main effects of AR image height ($F(3,75) = 29.2$, $p < .0001$, $\eta_p^2 = .339$) and real world-target distance were established ($F(2,114) = 21.3$, $p < .0001$, $\eta_p^2 = .272$).

Bonferroni corrected pairwise comparison analysis revealed reaction times were significantly faster when AR image placed above (2.4s) compared to when overlaid (2.8s) the real world target, ($p < .0001$). Further pairwise comparisons, using least significant difference corrections, discovered at closer distances reaction times are significantly faster (see Table 3.4); with the average response time being 2.2s at 10m, 2.6s at 20m and 2.9s at 25m. Reaction time analysis found no between subject effects, with the average reaction time for brightness (2.8s), size (2.5s) and size and brightness (2.5s) was not significant different to respond (see *Figure 3.10*).

Measure	Significance
10m x 20m	p < .00001
10m x 25m	p < .00001
20m x 25m	.029

Table 3.4: Pairwise comparison using least significant differences for reaction time analysis for differences between real-world target distances

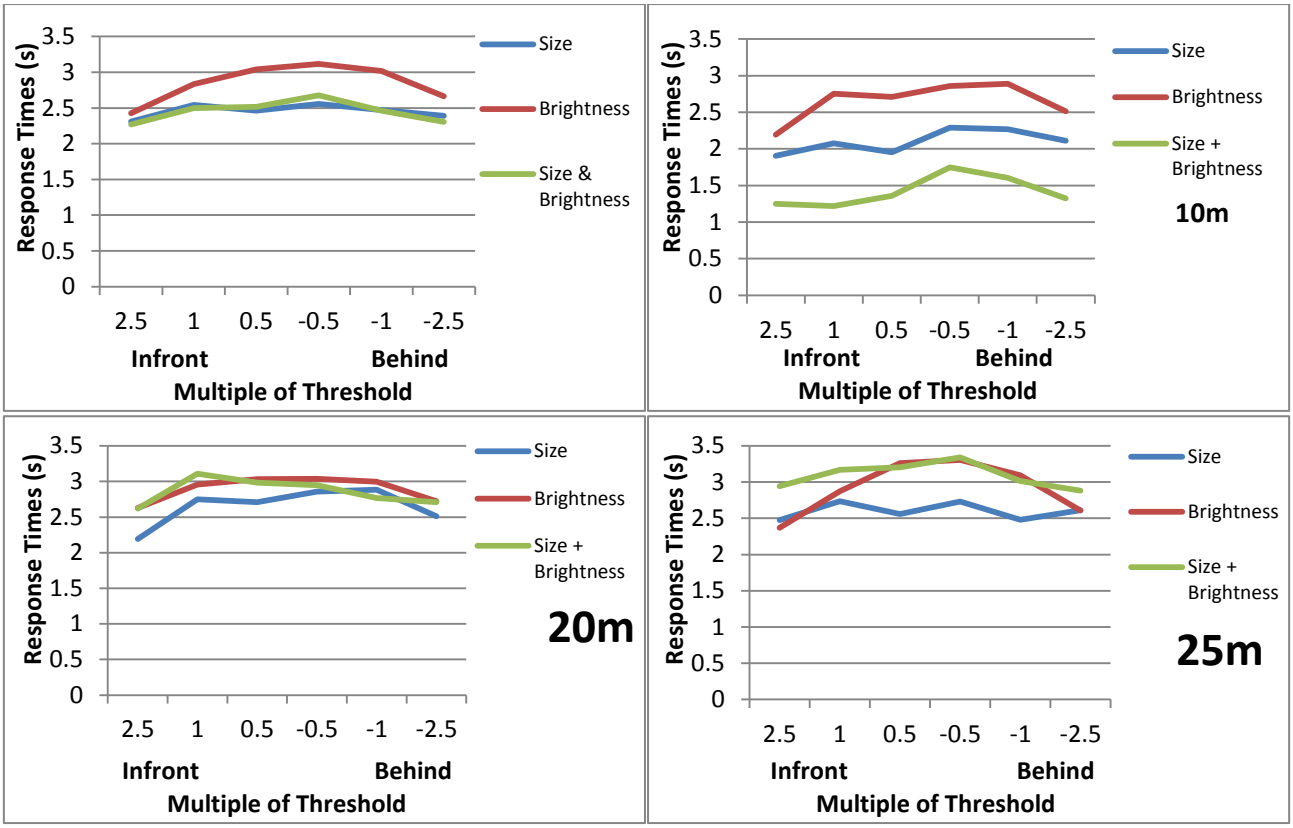


Figure 3.10: Reaction time graphs comparing the across the three depth cues; a) overall reaction times across all distances tested, b) reaction times for 10m, c) reaction times for 20m and d) reaction times for 25m.

3.8 Discussion

A set of three experiments explored the impact of relative depth cues for size and brightness across a range of distances in AR depth perception. By isolating and combining together cues of relative brightness and relative size, the investigation aimed to understand how depth cues combine in an AR environment. Findings from the experiments aim to develop theory on current and possible future models of depth perception in AR environments and devices.

3.8.1 Brightness not significantly inferior to size as a depth cue

The brightness cue condition was found to have the largest depth localisation error for participants, however brightness was not significantly different to size cue. When comparing slope analysis between size only and brightness only cue conditions, across all four test distances, size again was not statistical superior to brightness only condition. The finding that in AR depth perception size is not significantly superior to brightness has been demonstrated before by Surdick et al (1994) and Surdick et al (1997). However supporting findings were tested at only 1 and 2m.

Finally when comparing the current experimental results with the experiments in Section A (stereopsis) only size and brightness combined significantly improved depth perception performance. A finding which highlights the importance of stereopsis as a depth cue for AR depth perception of upto 25m, similar to real world depth perception (Allison et al, 2009).

3.8.2 How results are understood in terms of cue combination

When comparing results of the current experiment to the previous study of the chapter (Stereo only condition) only significant differences between 'stereo size and brightness' cue condition were established. Both 'stereo and brightness' and 'stereo and size' when analysed did not produce significantly superior depth performance from participants. Firstly this finding illustrates the importance of stereopsis in AR depth perception for upto 25m; replicating real-world results on far field stereopsis depth perception (Allision et al, 2009).

This finding may also indicate the importance of combining cues in AR. Only when cues of size and brightness were combined with stereo did observer depth perception significantly improve against just stereo. How does this apply to models of cue combination?

Firstly it can be assumed that the 'winner takes all' model does not account for the current results; as a combination of all cues (stereo, size and brightness) was significantly superior to 'stereo' only condition. The MLE would predict equal contribution of size and brightness cues having equal impact: this is a possibility as there was no significant difference established between size only and brightness only conditions.

The finding that the cue combination of stereo, size and brightness all combined together for improved depth judgments suggests a possible Bayes theorem application or the MWF. MWF proposes of a linear model for cue combination, as weights are based on reliability and availability; with an interaction of cues for cue promotion. Bayes theorem suggests computations are optimally combined with the mean and variance of cues representing the most likely stimulus and probability of availability.

3.8.3 Overlaid presentations produce less error than above

Exploring the direction of error found that in different cue conditions and virtual image locations results either produced underestimation of depth or an absence of underestimation. Analysis revealed for projections above the real-world target always produced an underestimation of depth; supporting a large body of previous work (Wither & Holler, 2005**).** However in both the size and the combination size and brightness conditions, when the virtual image was overlaid on the real world target depth judgement no underestimation was found; a finding which supports (Livingston et al, 2003**).** Overall estimations overlaid on the real world were superior to those presented above, specifically when the size cue was available.

3.8.4 Reaction times

When comparing reaction times between cue conditions, there were no significant differences established. This goes against previous work

demonstrating significantly faster reaction times for size only (Wither & Holler, 2005). The current result further demonstrates the difficulty in which participants are having in perceiving depth perception in AR.

3.8.5 Study Limitations

Although brightness not having much of an impact maybe because of how the stimuli was presented. Only an outline of a diamond, with some users reporting it was faint at very far away distances. Brightness cue may have had more of an impact as a solid shape, although for AR applications which aim to highlight real world targets, and avoid occluding with imagery, highlighting targets with an outline is the most likely way systems may use virtual imagery, so designers and engineers will need to apply size of imagery rather than brightness as a cue. This may be good for developers as systems already have issues with presenting the imagery and high brightness levels; such as vehicle HUDs will be able to use cues to size and occlusion of brightness may be a highly variable cue to AR depth.

3.8.6 Implications for AR devices

Overall findings reveal how cues to depth may differ in AR compared to real world depth perception, and how different cues may combine for users to have optimal depth perception in AR environments. Brightness has been found to be a poor indicator for depth perception in AR; with brightness only having an impact as a combination with size and occlusion. Previous work has often highlighted the impact perspective cues to depth have on AR (Surdick et al 1997, Wither and Holland, 2005); being described as possibly the best indicators for depth in AR (Surdick et al 1997), with peripheral information being more important potentially than the virtual image itself (Jones et al, 2008).

3.8.7 Summary & Conclusion

A study was carried out to assess how relative cues in depth perception interact in an AR environment. Relative size, brightness and a combination of both were compared at real world target distances between 5 – 25m under stereopsis viewing. Results revealed no significant differences for depth judgments between relative size and brightness, an unexpected finding which does not replicate what has been established in past psychophysical studies. Depth judgements improved on stereopsis only depth estimates when size and brightness combined together.

Results demonstrate how potentially depth information is acquired in AR, which is unlike that of real-world depth perception. It appears that the virtual image in an AR environment behaves that of a VR element; individual cues are unreliable. This leads to the question do users acquire depth information from the virtual object, the real world background, or are they combined in AR environment? The final section of the current Chapter presents an experiment which examines depth cues from the virtual image or the surrounding background.

Section C: How is depth information acquired in AR: the virtual image or the real-world background?

3.9 Introduction

Depth thresholds for AR environments are higher compared to real world environments; findings from the previous set of experiments in the chapter indicate it is not because of the method AR has been projected or measured. However depth perception in AR is inferior possibly due to the reliability of cues for a virtual image. For instance relative size was not found to be significantly superior to relative brightness, supporting previous work in AR (Sedgwick et al, 1994). This is unlike previous work in psychophysical research of depth perception. Results from the previous section indicate cues are effective when combined together. The final experimental work of the chapter examines if depth perception in AR is poor because information is derived from the unreliable virtual image compared to the more stable real-world background.

3.9.1 Virtual image or real-world background

Linear perspective cues in an AR environment are often a part of the natural real environment rather than the virtual image itself. Therefore it may be reasonable to assume a surrounding natural environment with more naturalist cues to depth may have more impact on depth perception than the depth cues to the virtual image. Applying depth cue theory, it would indicate the more cue rich an environment is the more accurate depth estimates would likely to be (Cutting, 1995).

Landy and Kojima (2001) exclaimed a potential reason for perceptual issues in AR is that they are cue poor environments. Compared to real world environments there are less depth cues; specifically from the virtual image itself.

When evaluating the contribution of linear perspective induced by real world objects in AR, the potential question asking ‘are observers acquiring depth information from the real world background or from the (potentially less reliable) virtual image?

Allison et al (2009) study examining intermediate distance depth perception in a real-world setting demonstrated depth constancy (perceived depth is proportional to the square of perceptual distance, (Allison et al, 2009)) was improved in conditions of rich natural cue conditions. A rich natural scene involved variety of objects such as chairs, desks, and computers. From this result the authors were able to raise the question; do humans acquire depth perception from the object being judged or from peripheral information derived from an environment?

In an augmented environment peripheral depth information may have a greater impact on distance judgements due to the variability of depth information derived from a virtual object.

Jones et al (2011) investigated peripheral visual information and its effects on medium field distance judgments in real world and AR environments. Implementing a blind walk paradigm (target stimulus being a wireframe pyramid), a between subjects design compared observers either viewing through a fully occluded AR HMD (preventing exposure to peripheral information for observers) or a real-world environment perceived through a HMD. Examining depth judgments between 3 – 7m, there was no significant difference for depth judgements between the occluded AR condition and the real-world viewed through a HMD condition. The researchers concluded from the finding that the majority of depth information acquired in an AR environment does not come from the virtual objects presented, however comes from the surrounding

environment (Jones et al, 2008). This conclusion would explain why underestimation of distance is higher in VR compared to AR.

Further examinations of the influence of peripheral information on AR depth perception have investigated with the use of augmented peripheral objects. Kyto et al (2011) examined stereoscopic depth judgements for virtual objects in an augmented reality environment; only upto 1.6m. In the experiment auxiliary augmentations, separate to the central target augmentation, were presented in a condition. Results revealed stereoscopic depth interval judgments were significantly superior with auxiliary augmentations presented in the scene. Also judgments were found to skew towards the user; a finding repeated so far in the experimental work of this chapter. The researchers concluded that the presence of the auxiliary augmentations aid users depth judgements in ambiguous scenes. In the research peripheral information was manipulated by virtual imagery.

It is unclear however from the above research (Jones et al, 2008; Kyto et al, 2011) in AR environments if peripheral information (the real world background scene) has more of an effect on depth perception on a virtual target, than the cue deriving from the virtual target at longer distances (upto 25m). In the current experiment rather than using a HMD and blind walking paradigms, an experiment will implement a perceptual matching task for precise depth thresholds without using an HMD. To date there has been no experimental assessment of linear perception manipulated by the real world scene examining distance perception at intermediate distances (< 30m).

3.9.2 Research Question & hypothesis

The goal of this study was to establish if in AR environments users attain depth information from the surrounding environment or the virtual image presented. This was done by creating two background settings, with one

inducing linear perspective, and the other reducing cue of linear perspective. Do users derived depth information from the virtual image target or more reliable surrounding real world depth information?

- Hypothesis 1: Reliable depth cues of real-world produce the best depth estimates in an AR environment
    - Prediction 1.1: Users will attain information of depth from surrounding real world background over the virtual image target.
        - Tested by: ANVOA interaction and post hoc analysis comparing increased real-world background conditions against increased virtual relative cues conditions.

3.10 Method

3.10.1 Participants

A total of 21 participants (10 male, 11 female) took part in the experiment. A mean age of 23.7 (SD = 2.8) with an age range of 19 – 31 years. All participants were either students or members of staff at the University of Nottingham. A stereo acuity TNO test was administered to each participant, to make sure all had normal depth perception.

3.10.2 Experimental design

For this study, participants would take part in four experimental conditions. Firstly there were two manipulations of the AR image. In one the participants would be able to use the size and brightness as cues to depth for the AR diamond, and in the other the diamond would be controlled, so relative size and brightness of the diamond could not be used to determine distance.

There were also two manipulations of the real world scene participants would be viewing. In one, increased depth information was produced with two

rows of traffic cones adjacent to the real-world target and a carpet (see figure 3.12). The other background scene removed the traffic cones and carpet.

With two manipulations of the AR image and two manipulations of the background scene there were 4 conditions each participant had to carry out: perspective with relative cues, perspective without relative cues, reduced perspective with relative cues, and reduced perspective without relative cues.



Figure 3.11: Augmented reality scene using traffic cones to induce depth cue of linear perspective

3.10.3 Procedure

Participants were given two full practice trials to become accustomed to the imagery. For each position in depth, for the AR image, there were three repeats of the task. Each block took around 5 minutes. After each block the pedestrian was moved to a new distance. Different virtual image positions were presented in different (counterbalanced) blocks, with a short break for the

participant in between. Each block of trials consisted of 18 diamonds. For the AR imagery positions, the distances for each block were completely randomized by the experimental software. On average each cue condition lasted 30mins; each participant carried out all four experimental conditions, with testing taking up to 2 hours.

3.11 Results

3.11.1 Real world background or virtual image cues

A three-way 2 (height: overlaid, above) X 3 (distance: 10, 20 and 25m) X 4 (Experiment condition: Increased background with virtual cues, increased background reduced virtual cues, decreased background with virtual cues, decreased background reduced virtual cues) within subjects ANOVA was completed. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(20) = .290$, $p = .014$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .67$). A significant three way interaction was established ($F(6,220) = 4.0$, $p = .001$, $\eta_p^2 = .167$), as well as main effects of height ($F(1,20) = 21.5$, $p < .0001$, $\eta_p^2 = .519$), distance ($F(2,40) = 79.1$, $p < .0001$, $\eta_p^2 = .798$) and experiment condition ($F(3,60) = 10.2$, $p < 0.0001$, $\eta_p^2 = .337$).

Examining the main effects of the three way ANOVA post hoc analysis using Bonferroni corrected pairwise comparisons revealed significant differences between 10m and 20m ($p = 0.42$), 10m and 25m ($p < .0001$) and 20m and 25m ($p < .0001$). Further exploring the main effect of experimental condition, pairwise comparisons found significant differences between 'increased background cues with virtual cues' and 'increased background reduced virtual cues' ($p = .008$), 'decreased background with virtual cues' ($p = .001$) and 'decreased background reduced virtual cues' ($p = .001$). These results

reveal that the viewing condition where both background cues and virtual cues were available produced the best depth judgments.

To further understand the three-way interaction and the differences between the four experimental conditions, specifically how the presence of either background cues or virtual cues impact of depth judgments in AR comparisons were made between experimental groups and virtual image height.

Comparing conditions of increased background with reduced virtual cues and decreased background with virtual cues a two-way 2 X (experimental group) 2 (height) within-subjects ANOVA was carried out. After assumptions of sphericity were reached, an interaction of F (2,80) = 8.1, $p < 0.0001$, $\eta_p^2 = .168$. Main effects of height (F(1,40) = 6.2, p = 0.01, $\eta_p^2 = .131$) and experimental condition were found (F(1,40) = 13.1, p <0.0001, $\eta_p^2 = .247$).

Post hoc analysis compared depth judgments between different heights and distances between the experimental groups. Six paired sample t-tests were carried out (corrected alpha was set to .008) and found significant differences between experimental conditions for overlaid positions at 25m (t (3.9), p = .001), and above position at 25m (t (3.4), p = .003). For the comparisons at 25m above and 20m overlaid superior depth judgments were found for increased background cue reduced virtual cue condition. For the overlaid at 25m comparison the difference was due to directional bias (see *Figure 3.12*).

This result indicates that participants make significantly superior depth judgments when only have real-world background cues to derive depth information when compared to only virtual cues at distance above 20m. Interestingly at overlaid positions approaching vista space there is a positional bias between conditions.

Comparing experimental conditions for background cues with increased virtual cues against condition of reduced background cues and virtual cues a two-way 2 X (experimental group) 2 X (height) within-subjects ANOVA was completed. After assumptions of sphericity were reached established no interaction or main effects (see *Figure 3.13*). This result suggests that depth information is not combined between the two sources of depth information and there may be a possible celling effect when combing cues to depth in AR.

Another comparison between groups compared increased background cues and increased virtual cues against increased background without virtual cues. A two-way 2 X (experimental group) 2 x (height) within-subjects ANOVA was completed. After assumptions of sphericity were reached established a significant interaction ($F(2,80) = 8.5$, $p < .0001$, $\eta_p^2 = .175$). No main effects were established.

Following the three-way interaction found between height X distance X experimental condition analysis examined possible interactions between height X distance across each experimental condition. First a two-way 2 (height) x 3 (distance) within-subjects ANOVA for 'increased background cues with virtual cues' condition, after assumptions of sphericity were met, established an interaction ($F(2,40) = 10.6$, $p < 0.0001$, $\eta_p^2 = .347$). No main effects were established.

A second two-way ANOVA for experiment condition of increased background and reduced virtual cues was carried out. A 2 (height) x 3 (distance) within-subjects ANOVA. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(2) = .618$, $p = .013$, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\varepsilon = .72$). Analysis found an interaction ($F(2,40) = 4.5$, $p = .017$, $\eta_p^2 = .185$) and a main effect of height ($F(1,20) = 6.8$, $p = .015$, $\eta_p^2 = 253$). Bonferroni corrected

pairwise comparison found projections overlaid were superior to those presented above (p = .03).

A third two-way 2 X 3 within-subjects ANOVA for experiment condition of 'decreased background with virtual cues' was carried out. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(2) = .715$, p =.049, therefore degrees of freedom were corrected using Huynh-Feldt estimates of sphericity (ε = .77). Analysis found an interaction ($F(2,40) = 3.6$, p = .035, η_p^2 = .154) and a main effect of distance (F (1,20) = 5.0, p =.011, η_p^2 = 202). Bonferroni corrected pairwise comparison found projections at 10m were significantly superior to those presented at 25m (.032).

A final two-way ANOVA for decreased background reduced virtual cues' was carried out. A 2 (height) x 3 (distance) within-subjects ANOVA established no interaction or any main effects. A Mauchly's test indicated that the assumption of sphericity had been violated: $X^2(2) = .001$, p <.0001, therefore degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity (ε = .50). A 2 (height) x 3 (distance) within-subjects ANOVA established no interaction or any main effects.

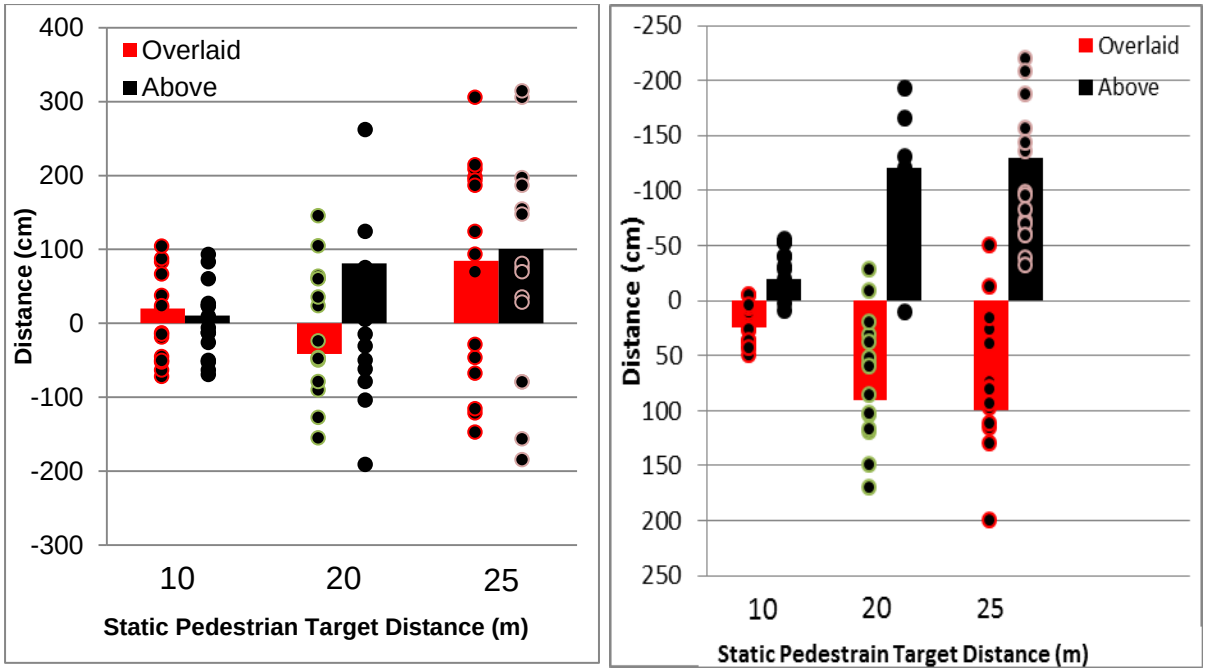


Figure 3.12: Depth localisation judgments for increased background cues/increased virtual image cues (left) compared to judgments for reduced background/increased virtual cues.

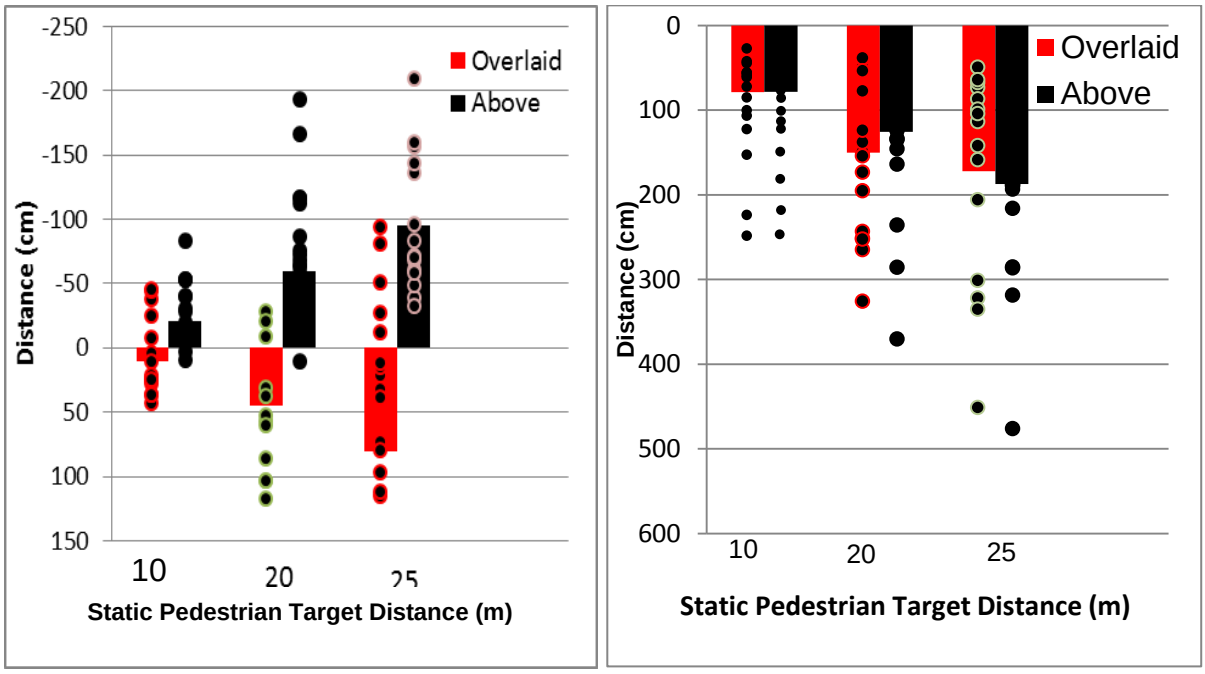


Figure 3.13: Depth localisation judgments for reduced background cues/increased virtual image cues (left) compared to judgments for reduced background/reduced virtual cues (right)

3.11.2 Slope analysis

To establish differences in directional bias for participant responses to the AR image between background cues and image cues, a 4 way within subjects ANOVA was carried out using the slope of the psychometric function. A 2 (Background: perspective, reduced perspective) x 2 (Virtual Image: relative cues, reduced relative cues) x 2 (Image height: overlaid, above) x 3 (Distance: 10m, 20m, and 25m) was used. After assumptions of sphericity were achieved no significant interaction was established. Main effects of real-world target distance ($F(2,40) = 52.7$, $p < .0001$, $\eta_p^2 = .725$), and virtual image cues ($F(1,20) = 13.8$, $p < .001$, $\eta_p^2 = .409$) was found, with Bonferroni corrected pairwise comparisons revealing added depth cues to virtual impact make a significant difference with size and brightness condition significantly superior to brightness only ($p < .0001$).

3.11.3 Reaction times

Participant's response to depth judgements was recorded and analysed using a four-way 2 (Background: perspective, reduced perspective) x 2 (Virtual Image: relative cues, reduced relative cues) x 2 (Image height: overlaid, above) x 3 (Distance: 10m, 20m, and 25m) was used. After assumptions of sphericity were met, there was no significant interaction found (see *Figure 3.14*). However a main effect of distance was established $F(4,10) = 2.7$, $p < 0.0001$, $\eta_p^2 = .251$). Bonferroni corrected pairwise comparison analysis revealed reaction times were significantly faster when AR image placed above compared to when overlaid the real world target, ($p < .0001$).

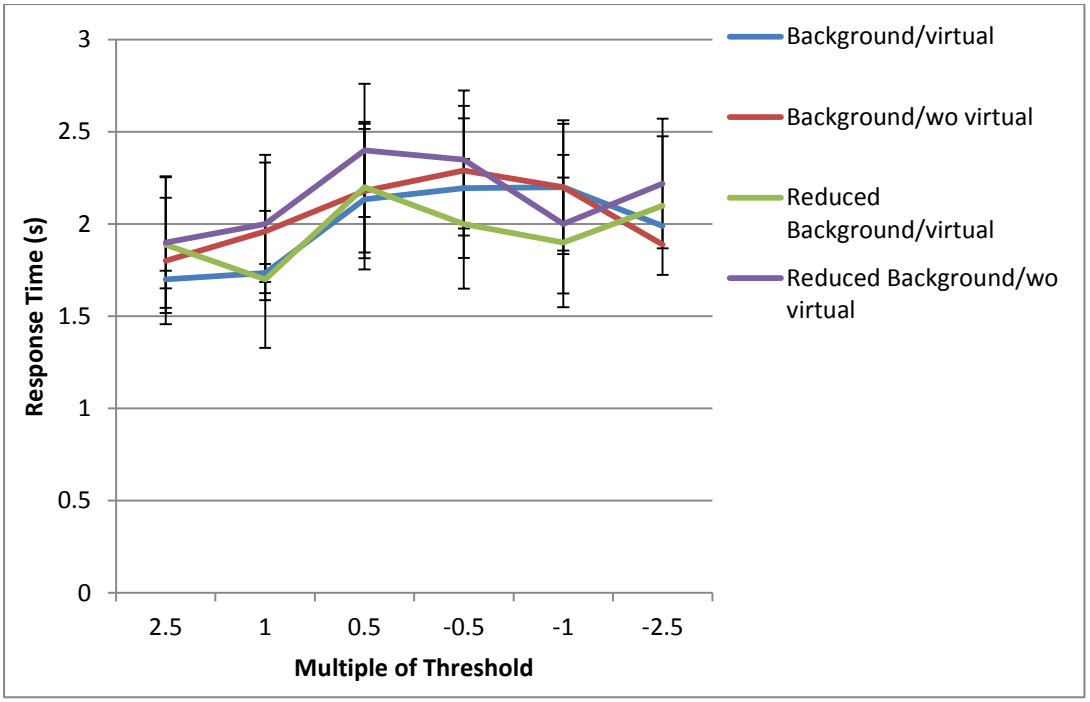


Figure 3.14: Average reaction time for all four experimental conditions. Error bars represent S.E.M

3.12 Discussion

In the final study of the chapter an experiment investigated where in a scene depth information is derived from in an AR environment; the virtual image or the real-world background. Manipulations of the real-world background scene and virtual image were compared against each other. Comparing experimental conditions it was revealed that the condition of background cues produced superior judgments compared to conditions only using increased relative virtual cues for judgments beyond 20m.

The result demonstrates that background information produces superior depth information for AR uses than the virtual image. A finding which supports Jones et al (2011) whom concluded that it is possible the majority of depth information is attained from the surrounding environment and not the object itself.

3.12.1 Depth information gained from either the surrounding world or the virtual image?

An interesting finding was that when background and virtual image cues were both available results did not produce improved depth perception. The conclusion could be highlighting a celling effect in AR depth perception. There becomes a point where additional information begins to flatten out in terms of performance. A similar finding was reported by Surdick et al (1997). However distances tested between studies hugely differ, as well as the stimuli used; Surdick et al (1997) used greyscale pixelated images, very different to the augmented green diamond used in the current set of experiments.

It is maybe more accurate to suggest in AR depth perception some users seem to either acquire information of depth from the virtual object target or from the real world background scene; participants seem to choose the depth cues which work for them.

This result may have a risen because participants are attaining the available depth cues of either the virtual image or the background; thus unable to combine them both for the environment. The mismatch in virtual cues and real world depth cues is still a novel experience to users; integrating both parts of the scene is a difficult task. However Jones et al, (2011) demonstrated users do become experienced with imagery and are able eliminate the 'novelty effect;' thus the mismatch in virtual and real depth cues should not be an issue.

Attention issues however, especially in experimenting, may cause users only to attend to information from either the real or virtual cues; such as a winner takes all model. This may be a potential experimental confound as testing took up to 2 hours, so any novelty effect worn off, may have caused attention deficits in processing.

Future work in AR will need to focus on attentional abilities of users to understand where exactly depth information is being derived from; the image or the background. Also the primary task was the judge the distance of the virtual image, therefore participants attention may have focussed on the depth cues of the image over the more reliable real-world background.

Another interesting finding for the study suggests there is a difference in positional bias between deriving depth information from background cues and virtual cues when the image is overlaid on to an object. When virtual cues are available it appears users begin to overestimate depth perception when imagery is overlaid at far distances. This may occur because the relative cues to depth are interacting with a possible occlusion effect derived from overlaying imagery directly over the real-world object.

If this is the case then it is of high importance for designers as for where in depth virtual imagery is positioned will depend greatly on if imagery is overlaid on a real-world target or not.

3.12.2 Summary & conclusion

The aim of the experiment was to understand if participants use depth cues of the virtual image or the real world background in estimation in AR environments?

Results revealed background information produces superior depth judgments compared to virtual image cues. There also seems to be a possible switch of overestimation of depth for when relative virtual image cues are available for uses when imagery is overlaid on the a real –world target. The current work supports past work finding an overestimation switch at 23m (Jones et al, 2008), and may help explain such a finding.

Results revealed that participants may not combine both virtual image cues and real world background. A finding showed that when both background and virtual cues as are available depth perception was not superior to a condition of where virtual cues where available. In the following and final chapter of the thesis a discussion will be presented on how participants in terms of previous cue combination models.

Chapter 4: General Discussion

4.1 Introduction

AR is being utilised into devices and systems for a variety of reasons. For instance AR technology is used for:

1) Devices which aid users when concurrently carrying out multiple tasks
2) Presenting virtual imagery on wide displays in users periphery to highlight real-world objects of interest
3) Highlighting real-world objects for specific locations in depth

The aim of the thesis was to apply psychological theories, concepts and experimental design procedures to the testing of dual-task attentional demands when processing imagery on wide field of view display and for measuring depth judgments in AR environments.

4.2 Wide field of view and attention

Psychology research examining attention using foveal and peripheral presentations has produced conclusions with regards to how humans process relevant and irrelevant stimuli; in the form of early and late attentional selection. Specifically perceptual load theory proposes that attentional selection is more selective under higher perceptual load; complexity of stimulus processing (Lavie, 1995; Lavie & Tsal, 1994). Past examinations have usually manipulated perceptual load by increasing the number of irrelevant distractors, presentation time of stimuli, task switching, effects on working memory (e.g. Lavie, 1995; Lavie et al, 2004). In this thesis the idea that peripheral target eccentricity from central presentation is a form of perceptual load is examined with regards to foveal task difficulty; as well as visibility of peripheral targets. Psychology research however has rarely examined perceptual load mechanics for wide FOV displays.

Human factors research has examined UFOV for large display areas (e.g Miura, 1990; Crundall et al, 1999, 2002; Van Weijgert, 1991, 1993; Seya et al, 2013). Conclusions have been split between cognitive tunnelling and general interference. Certain experimental manipulations such as the difficulty and presence of foveal task have been detrimental to the outcomes (Williams, 1988, 1989). However these findings map closely to perceptual load predictions with regards to the occurrence of perceptual load theory.

Combining the research methodologies of wide FOV studies in human factors and psychological theories of attention (specifically perceptual load theory) the thesis presented experiments manipulating peripheral target eccentricity and presentation time to help understand how these manipulations are impacted by the increasing demands of a foveal task.

4.2.1 FOV and attention conclusions in context of perceptual load theory

Over the course of two experiments assessing the impact of increased foveal task difficulty on peripheral performance revealed that attention allocation is more dependent on presentation time of peripheral stimuli than target position; in terms of increased eccentricity. This conclusion was demonstrated by two experimental findings. In the first experiment a significant two-way interaction between foveal task and peripheral task difficulty was established. Further post hoc analysis revealed peripheral performance decreased as a function of eccentricity. In the second experiment there was no interaction established between foveal task and peripheral task difficulty; post hoc analysis revealed that peripheral performance decreased as a whole rather than as function of eccentricity.

Past examinations of UFOV have often focussed on the nature of the foveal task, such as the number of investigations carried out for driving

simulations. The display of peripheral task presentation has rarely been tested beyond 30^{o} (Miura, 1990; Crundall et al, 1999; Seya et al, 2013; Gasper, 2016). Previous examinations have demonstrated only early selection of attention (Miura, 1990; Crundall et al, 1999; Seya et al, 2013), apart from Gasper et al, (2016) which demonstrated late selection, however did not manipulate the load of foveal task. In Section B a foveal task load was manipulated and a demonstration of late selection did occur; this may suggest further the importance of peripheral task on attentional selection.

However previous examinations demonstrated early attention when also presenting peripheral targets at similar time durations. For instance both Seya et al (2013) and Crundall et al (1999) presented peripheral targets for upto 200msec and both demonstrated tunnel vision. Difference between the current investigations and those previous used much small display areas; neither experiment used display areas larger than 15^{o}.

The increased presentation time of peripheral targets in Section B allowed participants to spread attention across both tasks more equally. This is represented by the increased performance on the peripheral task; without the cost to errors on the foveal task. The increased presentation time counters the increased demand of large display area and allowing for optimised allocation of attention; with attention alternating between the tasks.

Previous experiments have used the finding of an interaction between foveal cognitive load X peripheral target eccentricity as a demonstration of cognitive tunnelling; termed often as tunnel vision or selective interference (Williams, 1988).

As discussed before, the predictions for Chapter 2 applied perceptual load theory to past findings of UFOV. When a foveal task demands enough of a participant's attentional capacity, less of the peripheral visual field will be

processed. This will be represented by an increase in fixation difficulty and an omission of irrelevant peripheral targets at the early stage of processing; i.e. early attentional selection. This was demonstrated in the findings of the Section A of Chapter 2.

However in Section B the interaction between foveal load x peripheral target eccentricity was not established; this is often termed in UFOV research as a demonstration of general interference. Applying perceptual load theory to the result, it was concluded this was a demonstration of late attentional selection. As remaining attention was not focussed on the foveal task and was allocated to the peripheral task-irrelevant items.

This suggests that for a large display area irrelevant distractors are time-dependent rather than eccentricity dependent. This conclusion is made because there were only two experimental design changes for peripheral stimuli between Section A and Section B. The first was the peripheral target position and the second was the presentation time. If the result was dependent on peripheral target eccentricity a similar finding of early attention should have been demonstrated in Section B as in Section A as the display area was larger in Section B.

Another suggestion is that in Section B tunnel vision was not produced because the foveal task was not sufficiently difficult enough to demand enough attention from participant (Williams, 1988). However, the finding in Section B of tunnel vision (or early selection) demonstrates the foveal task did demand enough attention. There were no changes in foveal task between Section A and Section B for easy and hard conditions.

The current set of experiments demonstrates the load of peripheral task is a key factor in determining the occurrence of early or late selection; both in terms of display area and presentation time.

4.2.2 FOV and attention findings for design of AR devices

The experiment was designed to examine the capability of users to use AR displays with wide FOV; specifically vehicle HUDs. There are several concerns raised by the potential of a HUD encompassing an entire windscreen:

1) Are users able to accurately perceive visual information at far eccentricities?
2) Will peripheral target presentations at far eccentricities cause errors on the central fixation task?

4.2.2.1 Peripheral accuracy

Examining the results of Section A and Section B peripheral target performance suggests that users are able to accurately perceive stimuli in the periphery when targets are presented for 1 s. In Section A peripheral performance at 60^{o} averaged at 30% accuracy; this is at chance level. At 40^{o} presentations performance had only increased to 35% accuracy. Presentation time was only 25msec. These accuracy scores are low overall after averaging across contrast conditions. However even for the highest visibility conditions for peripheral targets (0.7), accuracy for presentations at 40^{o} and 60^{o} (both left and right directions) does not go beyond 40%. In Section B presentations at the farthest points of the display averaged at 75% accuracy for presentation times of 250msec.

4.2.2.2 Left and Right Visual Field

It is important to mention the observed left and right visual field difference observed across experiments. As discussed before this could be due

to physiological issues relating to the eye, or a behavioural development due to driving on the left hand side of the road. It is an important design constraint for AR HUD manufacturers. Designers need to take into account this deficit as it is the display area where the furthest peripheral targets will be presented.

It is unlikely the size of the stimuli was the cause of the low accuracy rates of peripheral targets as Landolt C's were presented at 18^{o}, (which is large). To increase accuracy rates presenting peripheral visuals for longer is a potential. However the duration at which peripheral stimuli can be presented depends on the type of information it is highlighting; an oncoming hazard, or point of interest demand different attention allocation.

A conclusion of the study relating to early and late attentional selection can inform designers about duration presentations of peripheral stimuli on a large AR HUD. To compensate for occurrences of early attentional selection, when peripheral performance decreases in the far periphery (when foveal difficulty increases) peripheral presentations could utilise visual imagery which activates bottom-up activation. For example presenting salient visual information in the far periphery of a windscreen to gain users attention

The increase in presentation time producing late-attentional selection indicates that large AR HUDS potentially removes distraction, rather than causing it. Therefore increased presentation time should only be helpful for users rather than distracting. An issue with early HUDs was that the small devices would result in cognitive tunnelling with drivers ignoring potential peripheral hazards (Gish & Staplin, 1995). Further concerns were raised for the possibility of increasing HUDs to encompass the entire windscreen, in that peripheral information would now distract users from the task of driving. However this has not been found in the current examination.

4.3 Depth perception

Previous research in AR environments has documented perceptual issues regards to depth perception in such settings; with experimental reports proposing a large underestimation of depth (Livingston, et al 2003; Kirkley, 2003; Wither & Holler, 2005; Swan II et al, 2006; Jones et al, 2008). However the exact amount of underestimation and error in depth has been debated, with experimental designs, methodology, settings and equipment used varying between studies and findings (Reiner, 2013).

Also due to the nature of presenting imagery in AR (the requirement for optical devices) it is difficult to directly compare studies in AR depth perception using HMDs and action based measuring designs to experiments examining real-world depth perception implementing perceptual matching studies.

In this thesis the aim was to examine AR in a similar manner to previous examinations of medium to far distance depth perception (e.g. Allison et al, 2009). This would allow for better understand to the extent of underestimation and error in depth judgments for AR compared to real-world environments. With additional analysis manipulating cues to depth for both virtual imagery and real-world backgrounds findings would help understand how depth perception in AR works. Does a real-world depth element behave like real-world environments and do virtual elements in AR behave in the same way as virtual reality?

Providing answers to these questions would develop understanding of psychological theory with regards to models of depth perception for environments where cues are possibly unreliable or unstable. As well as human factors design implications for AR HUDs aiming to place virtual imagery and the same perceived distance as the real-world target it is highlighting.

4.3.1 How does depth perception in AR compare to real-world

The result from the first experiment of Chapter 3 demonstrated that depth resolution is poor in AR compared to real-world. With results not dissimilar to previous experimentation in AR using perceptual matching designs (Swan II et al, 2006). Interestingly in the second study of the chapter it was demonstrated that relative cues to depth for virtual imagery in an AR scene have different impact of depth judgments to real-world depth perception. This is the first time virtual cues to depth have been measured and assessed for depth perception at far distances.

In Chapter 3 Section B, the depth cue of size was not found to be significantly superior to relative brightness. This result has been demonstrated before in AR (Surdick et al, 1994; Surdick et al, 1997) however only at short distances of 1 – 2m. However the current study tested at much further distances where the impact of size as a cue is greater compared to relative brightness, and differs to previous examinations of cues to depth (Cutting, 1995). Interestingly when the results of Section B were compared to Section A, neither size nor brightness were significantly superior to stereopsis depth perception; until relative cues were combined together.

It appears that cues in AR depth perception combine in a method to that of MWF. Relative cues of brightness and size appeared to combine together to make improved depth judgments. On their own it appeared relative cues are weaker (unreliable) for a virtual image and did not significantly contribute. It appears cues were weighted and combined together in an additive manner forming a robust combination rule which is based on a linear cue combination model (see *Figure 4.1*). It appears weights for relative cues to depth are dissimilar to real-world cues. This is also partially evidenced by Section C results which demonstrated that real-world background estimates produce

significantly superior depth judgments compared to virtual image. It also appears in AR depth perception there is a celling effect in AR depth perception (see *Figure 4.1*).

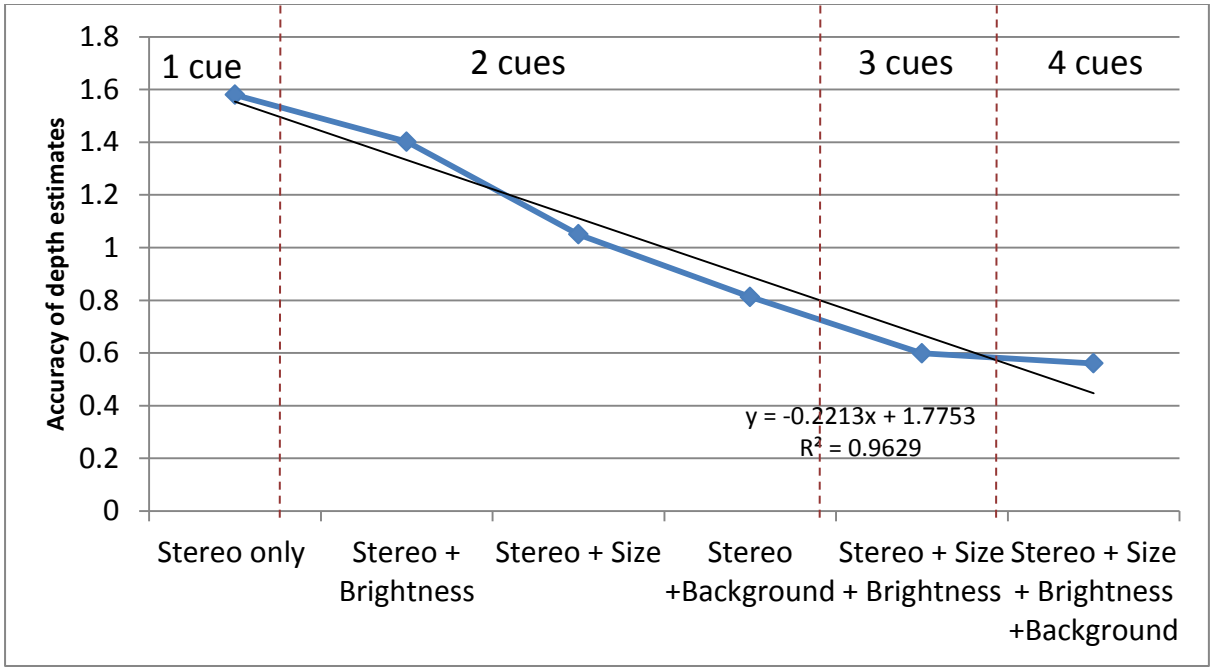


Figure 4.1: Graph representing the accuracy of depth estimates (calculated variance across participant scores) for different cue combinations. Some cues appear to have more of an impact than others, with an increase in accuracy with more cues. Appears however to be a celling effect

4.4 Conclusion

4.4.1 Contribution to knowledge

4.4.1.1 Attention and wide FOV

The current studies and findings are the first to explain and understand past findings of UFOV in terms of early and late selection and how peripheral load of targets impacts on attentional selection. Large viewing devices do not necessarily causes distraction to users as attention to distractors may be time-dependent rather than caused by increased peripheral eccentricity.

4.4.1.2 Depth perception in AR

With regards to depth perception the current work gives a clearer indication of what the capabilities of depth perception are in AR. Results give a clearer indication of exactly where perceptual equality for a virtual image in an AR environment is (for a variety of cues and cue combinations) how relative cues combine in AR is similar to a MWF model, and the impact background information has in AR.

4.4.2 Limitations of studies

The FOV and attention studies did not fully test or explore cognitive capabilities of a user. Many past examinations have tested working memory as well as purely visual constraints of attention. This would have possibly produced a fuller explanation of attentional abilities for UFOV. Certain visual properties of peripheral targets were not measured or assessed such as size of the targets. There were too few peripheral target eccentricities tested. For instance smaller targets at more positions could have produced regression analysis which showed the point in eccentricity where 75% accuracy for peripheral discrimination.

In the depth experiments the real-world target was a flat cardboard cut-out. In the real world real world objects are often not flat, but have depth to them. It is possible the flat object may have had an impact on judgments for overlaid presentations. Also the study did not examine monocular depth thresholds, which for a full model of depth perception is required.

4.4.3 Future directions

4.4.3.1 Switching attention between depth planes

3D space inhabits both depth in one direction and field of view in the other direction. The current examinations studied depth and field of view in separate experimental paradigms. However visual search requires attention to be

deployed to a precise distance and direction in a 3D environment. Only a limited amount of work has investigated the deployment of attention in 3d space.

Previous research has explored the effects of deploying attention across different depth planes (Finlayson & Grove, 2015; Roberts, Allen, Dent & Humphreys, 2015). For instance Finlayson & Grove (2015) measured the effects of 3d volume space, distance in depth, number of depth planes and relative position have on 3d visual space search for a real world environment. The participants' task was to search for a forward tilting bar amongst a search array of distractors (backward tilting bar). Set sizes consisted of either 12 or 24 items and distributed across two or four depth planes. The four stereoscopic depth plane magnitudes were 18 arc min crossed disparity (depth plane 1), 6 arc min crossed disparity (depth plane 2), 6 arc min uncrossed disparity (depth plane 3), and 18 arc min uncrossed disparity (depth plane 4), relative to the plane of the display. The initial position of the participants' gaze and attention was controlled with a Landolt C task. Results found a difference in response times to detect targets in each of the 3D display conditions, with findings indicating attentional deployment is impacted by 3D spatial layout (Finlayson & Grove, 2015). It was found that visual search performance depends on 3d volume, relative target position in depth, and number of depth planes. Interestingly an asymmetrical preference for targets in front of a display was found, demonstrating that arranging items in additional depth planes increases search time and search inefficiencies (Finlayson & Grove, 2015).

It is understood that attention can be deployed at specific depths (Downing & Pinker, 1985; Marrara & Moore, 2000) however there are a number of parameters which can influence an observer's ability to attend to precise depths, and the ability to switch in between depths (Finlayson & Grove,

2015). There is very little understanding in how an observer would switch attention from a real world object to a virtual object, or vice versa. As well how would parameters such as 3d volume space, distance in depth, number of depth planes and relative position have more or less of an influence in an AR environment? Is it more difficult (less efficient) to switch attention between real and virtual objects compared to switching attention between two real world objects? Investigating attentional shifts in depth would help develop an encompassing framework for how humans deploy attention across 3d space, and also aid designers for 3d displays.

Modifying the paradigm used in the current thesis to assess depth perception in an AR environment, it is possible to design a set of experiments which would examine how observers would attend to multiple depths in an AR environment. For instance, the paradigm has the potential to use two separate moving projector units to produce two separate virtual images at different depths. Testing will be able to measure reactions times and accuracy for selecting one of multiple targets in a scene (real and virtual), and the potential for switching between real and virtual objects.

4.4.3.2 Recommendations for practice (Vehicle manufactures)

The practical implications for the current research have been guided towards vehicle HUDs. It is seen as essential for vehicle manufactures to produce a set of guidelines for future technological applications (Hasedzic et al, 2015). The aims of AR HUDs include reducing eyes-off-the road time by presenting important information to a driver which does not cause distraction (safety). This information needs to be presented in a manner which is eligible and clear (quality) and when highlighting real-world objects in depth it needs to match in terms of perceptual equality for each object each time (reliable).

Imagery presented on AR HUDs will also need to take into account the surrounding conditions of the outside environment (normal and abnormal conditions).

With regards to HUD design the aim of testing was to begin to establish a potential boundary for design space for AR in terms of FOV and the depth resolution for virtual imagery in conjunction with real-world objects. Conditions of workload were used to represent different driving conditions to understand how a design space for AR may change in normal and abnormal conditions. The recommendations for practice have been split into three groupings: safety, reliability, and control under normal and abnormal conditions.

4.4.3.2.1 Safety

One of the initial objectives of the study was to develop a potential boundary for design space for an AR HUD. If an AR HUD encompassed an entire windscreen would the imagery presented in the far periphery distract a driver? Would a driver even be able to detect imagery? The experiments in Chapter 1 assessed the impact peripheral targets have on a central in-lane driving like task.

In both experiment 1 (horizontal eccentricities) and experiment 2 (vertical eccentricities) results found participant's ability to protect the CTT over peripheral target presented at the farthest eccentricities of the windscreen. The overall finding of central task protection begins to suggest that an AR HUD has the potential for encompassing the whole windscreen without causing distraction to the primary task of driving. However there are certain conditions where the size of the AR HUD may need to be reduced.

4.4.3.2.2 Normal and abnormal conditions

One of the advantages of AR HUDs is to present imagery to aid in driving for difficult conditions, such as low visibility (fog, heavy rain, and night). The current assessment did not investigate the low visibility of the external scene (only peripheral targets); however this is an important factor to test in the future as it will impact how virtual imagery may be presented. However the current work did assess the impact of increased central load on peripheral discrimination performance. In experiment 1 it was found that increased central load (potentially difficult driving conditions) resulted in finding of tunnel vision; a shrinking in a user's FOV as a function of eccentricity.

In experiment 2, after peripheral presentation time was increased, as central load increased there was a general drop in peripheral performance which was not a function of eccentricity. The change in presentation time was to represent different types of information potentially presented on an AR HUD. Short presentation time represented hazards requiring immediate allocation of resources, and increased presentation time for imagery such as navigation.

From these sets of results a suggestion for vehicle manufactures would be to potentially have a shifting size of presentation space. An AR HUD may need the capability to alter its size/space depending on the condition of driving. In difficult high workload driving conditions (heavy rain, heavy traffic, and city center rush hour) the presentation space may need to shrink in terms of available FOV, therefore reducing the chance of drivers missing important information in the periphery, especially in the case for hazards.

4.4.3.2.3 Reliability & Quality

Throughout the depth experiments conducted for the thesis, results found a large difference between participants. For instance where one participant may

see a virtual image match an external object when the image is over 2m infront, whereas another may perceive the image a match when placed 50cm behind the real-world object. This is a finding which has been frequently reported in AR depth perception (Swan II et al, 2008). For a driver to be able to reliably connect a virtual image in depth to the real-world target object it is highlighting it is strongly recommend that future AR HUDs are individually calibrated for each driver.

Users of AR HUDs will need to be able to reliably perceive a virtual object at the same distance as the real-world object it is highlighting. Not only do individual differences between users affect perceptual equality, but the research conducted in the thesis also indicates different cues to depth, and how they combine with each other, vary in reliability (see section 3.7). The knowledge of perceptual equality derived from different cues of depth is useful for designers in developing AR systems which feel more immersive, reduce potential visual complaints and produce faster more accurate reactions to virtual imagery presented.

Over the course of chapter 3 a total of six experiments were carried out applying different manipulations to both virtual imagery and real-world background. Data was collected using perceptual matching tasks; which required participants to make a judgment for if they perceived a virtual image either to be in front or behind the real-world object.

Calculations were used for the points in distance were participants perceived the virtual image to be in front (75% response in front), to be behind (25% response in front) and at the same perceptual difference (50% response in front) as a real-world object it is highlighting. Figures 4.2, 4.3 4.4 and 4.5 represent the points in distance for different cue combinations where a virtual image is perceived in front, behind and matches an real-world object.

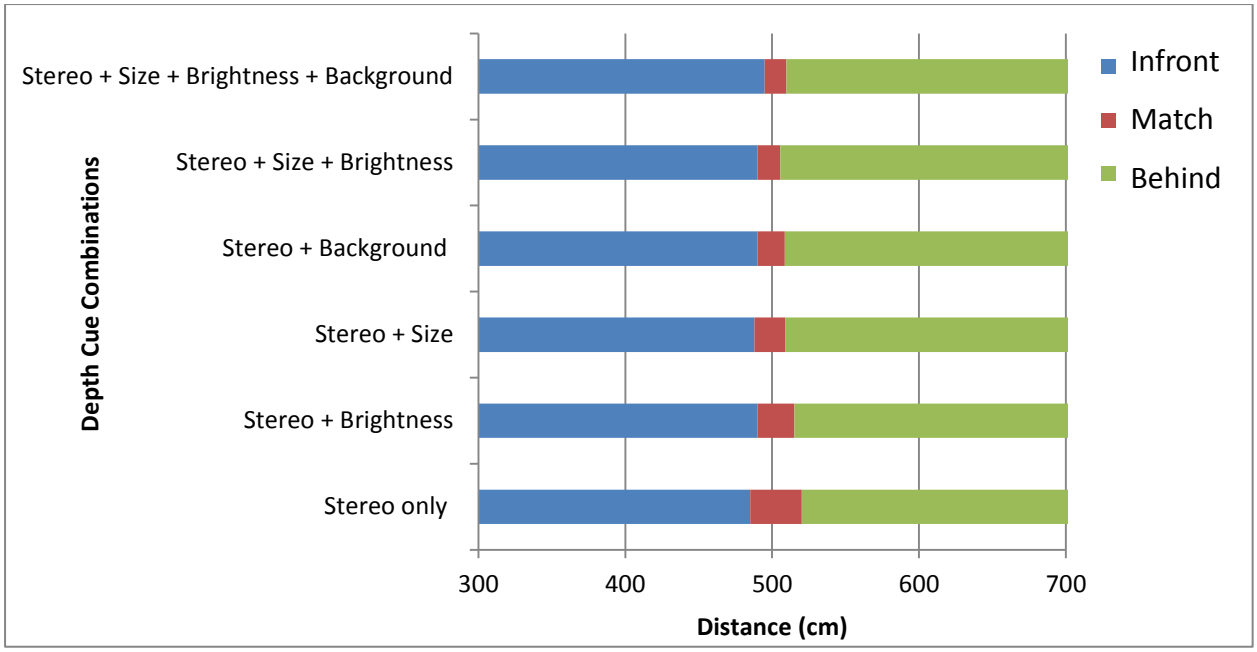


Figure 4.2: A graph representing the points in distance in which a virtual image is perceived as infront (blue), matching (red) and behind (green) of a real-world object at 5m for a variety of cue combination conditions

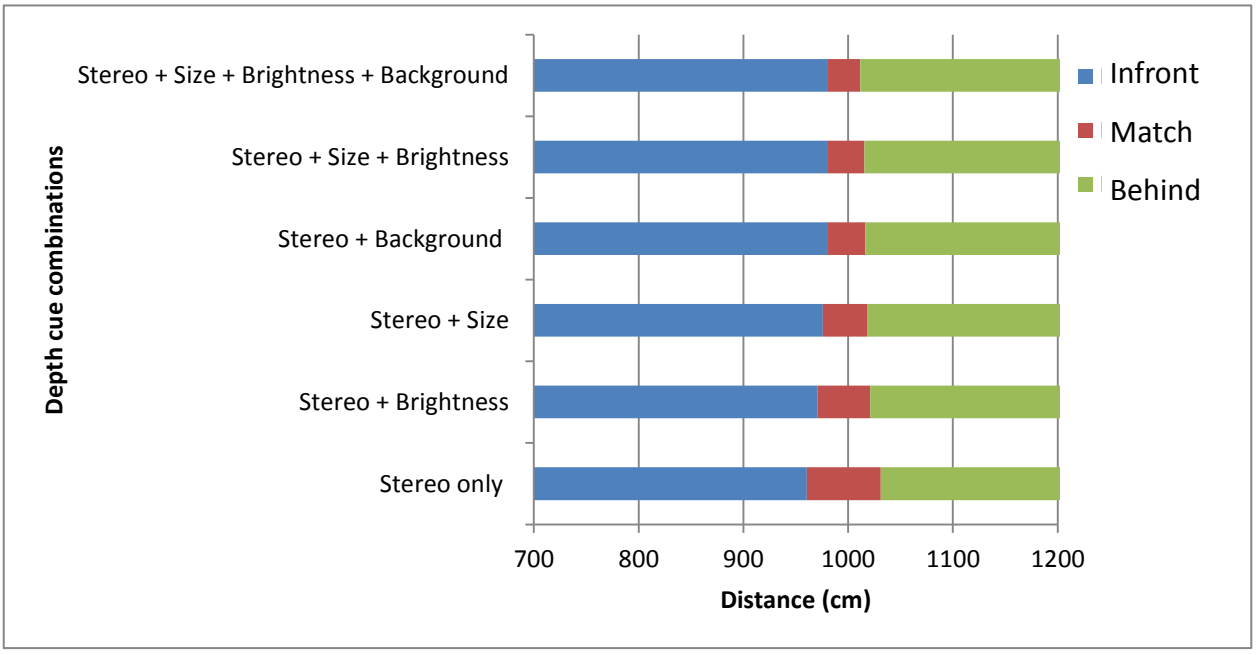


Figure 4.3: A graph representing the points in distance in which a virtual image is perceived as infront (blue), matching (red) and behind (green) of a real-world object at 10m for a variety of cue combination conditions

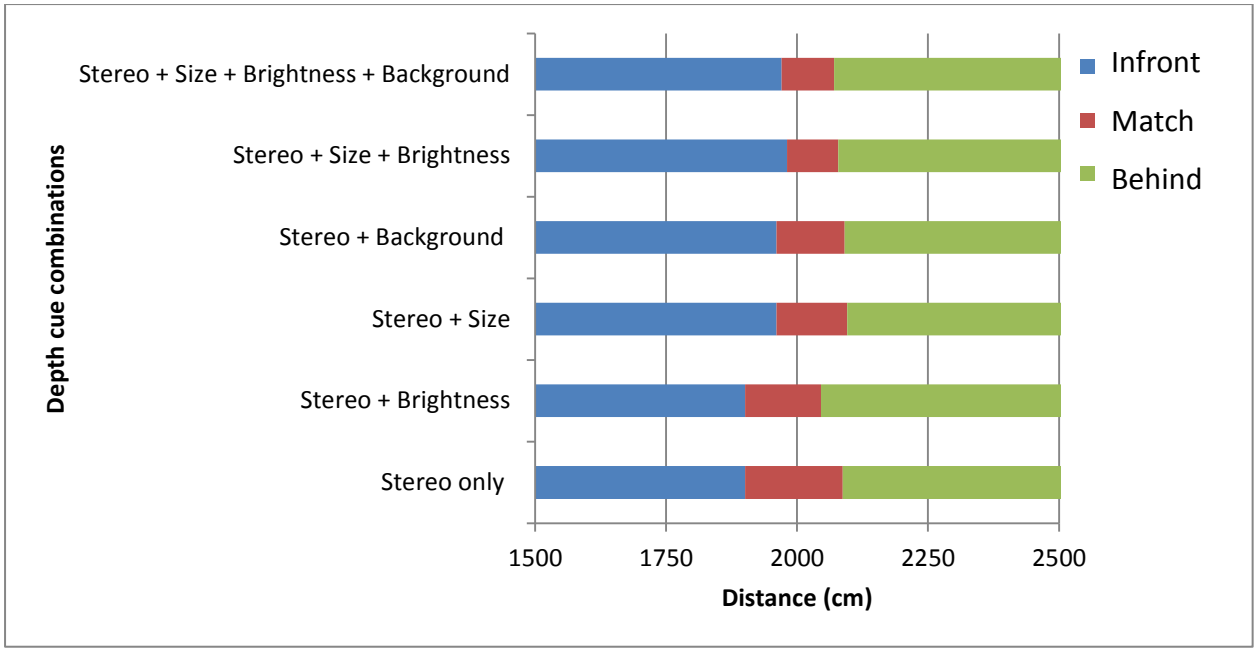


Figure 4.4 A graph representing the points in distance in which a virtual image is perceived as infront (blue), matching (red) and behind (green) of a real-world object at 20m for a variety of cue combination conditions

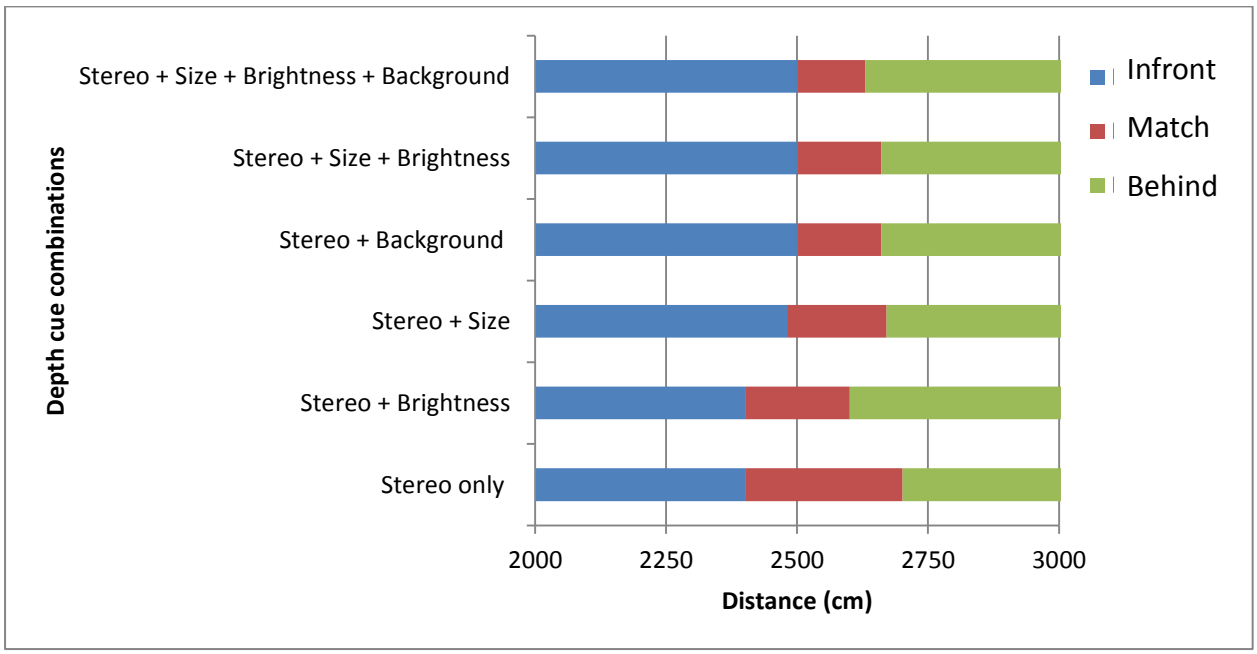


Figure 4.5: A graph representing the points in distance in which a virtual image is perceived as infront (blue), matching (red) and behind (green) of a real-world object at 25m for a variety of cue combination conditions

In the graphs the red area represents the area in space where a virtual image can be placed in distance and a user will be able to perceive it as matching the real-world target in terms of depth. As more cues to depth combine the smaller the space is. This means designers can create imagery which produces more precise and reliable depth estimates for users depending on the available cues to depth.

References

Abramov, I., Gordon, J., & Chan, H. (1991). Color appearance in the peripheral retina: effects of stimulus size. *JOSA A, 8*(2), 404-414.

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of vision, 9*(1), 10-10.

Angelaki, D. E., Gu, Y., & DeAngelis, G. C. (2009). Multisensory integration: psychophysics, neurophysiology, and computation. *Current Opinion in Neurobiology*, *19*, 452–458.

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, *14*(7), 589-592.

Armbrüster, C., Wolter, M., Kuhlen, T., Spijkers, W., & Fimm, B. (2008). Depth perception in virtual reality: distance estimations in peri-and extrapersonal space. *Cyberpsychology & Behavior*, *11*(1), 9-15.

Arthur, K. W., & Brooks Jr, F. P. (2000). *Effects of field of view on performance with head-mounted displays.* University of North Carolina at Chapel Hill,

B. Sekuler, P. J. B., Mortimer Mamelak, Allison. (2000). Effects of aging on the useful field of view. *Experimental aging research*, *26*(2), 103-120.

Ball, K., & Owsley, C. (2000). Increasing mobility and reducing accidents of older drivers. In: *Mobility and Transportation in the elderly*

Ball, K., Owsley, C., & Beard, B. (1990). Clinical visual perimetry underestimates peripheral field problems in older adults. *Clinical Vision Sciences*, *5*(2), 113-125.

Ball, K., Owsley, C., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1993). Visual attention problems as a predictor of vehicle crashes in older drivers. *Investigative ophthalmology & visual science*, *34*(11), 3110-3123.

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *JOSA A*, *5*(12), 2210-2219.

Ball, K. K., Wadley, V. G., & Edwards, J. D. (2002). Advances in technology used to assess and retrain older drivers. *Gerontechnology*, *1*(4), 251-261.

Barlow, H. B. (1957). *Journal of Physiology*. (London). 136:469.

Baylor, D., Lamb, T., & Yau, K.-W. (1979). Responses of retinal rods to single photons. *The Journal of physiology*, *288*(1), 613-634.

Beck, D. M., & Lavie, N. (2005). Look here but ignore what you see: effects of distractors at fixation. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(3), 592.

Besharse, J., & Bok, D. (2011). *The retina and its disorders*: Academic Press.

Blake, R., & Hirsch, H. (1975). Deficits in binocular depth perception in cats after alternating monocular deprivation. *Science, 190*(4219), 1114-1116.

Broadbent, D. E. (1958). The effects of noise on behaviour.

Burgess, D. G. (1990). High level decision efficiencies. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 431–440). Cambridge, MA: Cambridge University Press

Campbell, F., & Green, D. (1965). Optical and retinal factors affecting visual resolution. *The Journal of physiology, 181*(3), 576-593.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research, 51*(13), 1484-1525.

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Attention, Perception, & Psychophysics, 57*(8), 1241-1261.

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research, 37*(1), 63-82.

Carrasco, M., Talgar, C. P., & Cameron, E. L. (2001). Characterizing visual performance fields: Effects of transient covert attention, spatial frequency, eccentricity, task and set size. *Spatial vision, 15*(1), 61-75.

Carrasco, M., Williams, P. E., & Yeshurun, Y. (2002). Covert attention increases spatial resolution with or without masks: Support for signal enhancement. *Journal of vision, 2*(6), 4-4.

Cartwright-Finch, U., & Lavie, N. (2007). The role of perceptual load in inattentional blindness. *Cognition, 102*(3), 321-340.

Chan, H., & Courtney, A. J. (1998). Stimulus size scaling and foveal load as determinants of peripheral target detection. *Ergonomics, 41*(10), 1433-1452.

Charisis, V., Vlachos, G., Chan, W., & Arafat, S. (2010). *Early notification warning system for prototype head-up display: development and evaluation of traffic congestion and sharp turn warnings* (0148-7191). Retrieved from

Charissis, V., Naef, M., Arafat, S., & Vlachos, G. (2008). *On the Impact of User's Computer Knowledge on Driving Simulation Test Results-HUD Simulation Case Study.* Paper presented at the Proceedings of the European Annual Conference on Human Decision-Making and Manual Control, EAM.

Charissis, V., Naef, M., Papanastasiou, S., & Patera, M. (2007). Designing a direct manipulation HUD interface for in-vehicle infotainment. *Human-Computer Interaction. Interaction Platforms and Techniques*, 551-559.

Charissis, V., Naef, M., & Patera, M. (2007). *Calibration requirements of an automotive HUD Interface using a Virtual Environment: Methodology and Implementation.* Paper presented at the Proceedings of: International Conference in Graphics and Visualisation in Engineering,(GVE'07), Clearwater, Florida, USA.

Clark, S. E. (1969). Retrieval of color information from preperceptual memory. *Journal of Experimental Psychology, 82*(2), 263.

Corbett, J. E., & Carrasco, M. (2011). Visual performance fields: Frames of reference. *PLoS One, 6*(9), e24470.

Correa, Á., Lupiáñez, J., Milliken, B., & Tudela, P. (2004). Endogenous temporal orienting of attention in detection and discrimination tasks. *Attention, Perception, & Psychophysics, 66*(2), 264-278.

Croner, L. J., Purpura, K. & Kaplan, E. (1993). *Proceedings of National Academy of Science. USA* 90, 8128–8130.

Crundall, D., Underwood, G., & Chapman, P. (1999). Driving experience and the functional field of view. *Perception, 28*(9), 1075-1087.

Curcio, C. A., Sloan Jr, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: individual variability and radial asymmetry. *Science, 236*, 579-583.

Cutting, J. E. (1997). How the eye measures reality and virtual reality. *Behavior Research Methods, 29*(1), 27-36.

Cutting, J. E., & Vishton, P. (1995). Potency, and contextual use of different information about depth. *Perception of space and motion, 69*.

Daitch, J., & Green, D. (1969). Contrast sensitivity of the human peripheral retina. *Vision Research, 9*(8), 947-952.

Daniel, P., & Whitteridge, D. (1961). The representation of the visual field on the cerebral cortex in monkeys. *The Journal of physiology, 159*(2), 203-221.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *The Journal of physiology, 357*(1), 241-265.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827-1837.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological review, 70*(1), 80.

Downing, C. J. & Pinker, S. (1985). The spatial structure of visual attention. In: *Attention and performance XI,* ed. M. I. Posner & O. S. M. Marin. Erlbaum.

Edwards, J. D., Wadley, V. G., Myers, R. S., Roenker, D. L., Cissell, G. M., & Ball, K. K. (2002). Transfer of a speed of processing intervention to near and far cognitive functions. *Gerontology, 48*(5), 329-340.

Edwards, W., Lindman, H., & Savage, L. J. (1963). Bayesian statistical inference for psychological research. *Psychological review, 70*(3), 193.

Ellis, S. R., & Menges, B. M. (1998). Localization of virtual objects in the near visual field. *Human Factors, 40*(3), 415-431.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Attention, Perception, & Psychophysics, 16*(1), 143-149.

Fan, J., Byrne, J., Worden, M. S., Guise, K. G., McCandliss, B. D., Fossella, J., & Posner, M. I. (2007). The relation of brain oscillations to attentional networks. *Journal of Neuroscience, 27*(23), 6197-6206.

Fancher, R. E. (1996). *Pioneers of psychology*: WW Norton & Co.

Fechner, G. (1860). Elements of psychophysics. Translation, H. *Adler. Brietkoph & Härtel, 1*.

Finlayson, N. J., & Grove, P. M. (2015). Visual search is influenced by 3D spatial layout. Attention, Perception, & Psychophysics, 1–9.

Foley, J. M. (1980). Binocular distance perception. *Psychological review, 87*(5), 411.

Gabbard, J. L., Fitch, G. M., & Kim, H. (2014). Behind the Glass: Driver challenges and opportunities for AR automotive applications. *Proceedings of the IEEE, 102*(2), 124-136.

Gaspar, J. G., Ward, N., Neider, M. B., Crowell, J., Carbonari, R., Kaczmarski, H., . . . Loschky, L. C. (2016). Measuring the useful field of view during simulated driving with gaze-contingent displays. *Human Factors, 58*(4), 630-641.

Gibson, J. J. (1950). The perception of the visual world.

Gish, K. W., & Staplin, L. (1995). Human factors aspects of using head up displays in automobiles: A review of the literature.

Glickstein, M., & Rizzolatti, G. (1984). Francesco Gennari and the structure of the cerebral cortex. *Trends in Neurosciences, 7*(12), 464-467.

Grechkin, T. Y., Nguyen, T. D., Plumert, J. M., Cremer, J. F., & Kearney, J. K. (2010). How does presentation method and measurement protocol affect distance estimation in real and virtual environments? *ACM Transactions on Applied Perception (TAP), 7*(4), 26.

Groeger, J. A. (2000). *Understanding driving: Applying cognitive psychology to a complex everyday task*: Psychology Press.

Grossberg, S. (1982). Contour enhancement, short term memory, and constancies in reverberating neural networks. In *Studies of mind and brain* (pp. 332-378): Springer.

Gruber, N., Müri, R. M., Mosimann, U. P., Bieri, R., Aeschimann, A., Zito, G. A., . . . Nef, T. (2014). Effects of age and eccentricity on visual target detection. *Frontiers in aging neuroscience, 5*, 101.

Hecht, S., Shlaer, S., & Pirenne, M. H. (1942). Energy, quanta, and vision. *The Journal of general physiology, 25*(6), 819-840.

Helmholtz, H. v. (1855). Ueber die accommodation des auges. *Graefe's Archive for Clinical and Experimental Ophthalmology, 1*(2), 1-74.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual review of psychology, 50*(1), 243-271.

Hills, B. L. (1980). Vision, visibility, and perception in driving. *Perception, 9*(2), 183-216.

Holmes, D. L., Cohen, K. M., Haith, M. M., & Morrison, F. J. (1977). Peripheral visual processing. *Attention, Perception, & Psychophysics, 22*(6), 571-577.

Howard, C. H. J. (1919). A test for the judgment of distance. *American Journal of ophthalmology, 2*(9), 656-675.

Hubona, G. S., Wheeler, P. N., Shirah, G. W., & Brandt, M. (1999). The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. *ACM Transactions on Computer-Human Interaction (TOCHI), 6*(3), 214-242.

Ikeda, M., & Takeuchi, T. (1975). Influence of foveal load on the functional visual field. *Attention, Perception, & Psychophysics, 18*(4), 255-260.

Jameson, D., & Hurvich, L. M. (1959). Perceived color and its dependence on focal, surrounding, and preceding stimulus variables. *JOSA, 49*(9), 890-898.

Jerome, C., & Witmer, B. (2005). *The perception and estimation of egocentric distance in real and augmented reality environments.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Jerome, C. J. (2007). *Effects of Spatial and Non-Spatial Multi-Modal Cues on Orienting of Visual-Spatial Attention in an Augmented Environment*. Retrieved from

Johansen, S., & Juselius, K. (1994). Identification of the long-run and the short-run structure an application to the ISLM model. *Journal of Econometrics, 63*(1), 7-36.

Jones, J. A., Swan II, J. E., Singh, G., & Ellis, S. R. (2011). *Peripheral visual information and its effect on distance judgments in virtual and augmented environments.* Paper presented at the Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization.

Jones, J. A., Swan II, J. E., Singh, G., Kolstad, E., & Ellis, S. R. (2008). *The effects of virtual reality, augmented reality, and motion parallax on egocentric depth perception.* Paper presented at the Proceedings of the 5th symposium on Applied perception in graphics and visualization.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. *Varieties of attention, 1*, 29-61.

Kaufman, L. (1974). *Sight and mind: An introduction to visual perception*: Oxford U. Press.

Kaufman, L., Kaufman, J. H., Noble, R., Edlund, S., Bai, S., & King, T. (2006). Perceptual distance and the constancy of size and stereoscopic depth. *Spatial vision, 19*(5), 439-457.

Kaye, S. B., Siddiqui, A., Ward, A., Noonan, C., Fisher, A. C., Green, J. R., . . . Watt, P. (1999). Monocular and binocular depth discrimination thresholds. *Optometry and vision science: official publication of the American Academy of Optometry, 76*(11), 770-782.

Kersten, D., & Yuille, A. (2003). Bayesian models of object perception. *Current opinion in neurobiology, 13*(2), 150-158.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55*(4), 352.

Kirkley Jr, S. E. H. (2003). Augmented reality performance assessment battery (ARPAB): Object recognition, distance estimation and size estimation using optical see-through head-worn displays.

Knapp, J. M. (2001). *The visual perception of egocentric distance in virtual environments.* ProQuest Information & Learning,

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 and Virtual Environments, 13*(5), 572-577.

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43*(24), 2539-2558.

Kruijff, E., Swan, J. E., & Feiner, S. (2010). *Perceptual issues in augmented reality revisited.* Paper presented at the Mixed and Augmented Reality (ISMAR), 2010 9th IEEE International Symposium on.

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), 6.

Landy, M. S., & Kojima, H. (2001). Ideal cue combination for localizing texture-defined edges. *JOSA A, 18*(9), 2307-2320.

Landy, M. S., & Maloney, L. T. (2000). Combination of cues and priors in depth perception. *Perception ECVP abstract, 29*, 0-0.

Lansdown, T. C., Brook-Carter, N., & Kersloot, T. (2002). Primary task disruption from multiple in-vehicle systems. *ITS Journal-Intelligent Transportation Systems Journal, 7*(2), 151-168.

Larish, I., & Wickens, C. (1991). *Attention and HUDs: Flying in the dark.* Paper presented at the Society for Information Display International Symposium Digest of Technical Papers XXII.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences, 9*(2), 75-82.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science, 19*(3), 143-148.

Lavie, N., Beck, D. M., & Konstantinou, N. (2014). Blinded by the load: attention, awareness and the role of perceptual load. *Phil. Trans. R. Soc. B, 369*(1641), 20130205.

Lavie, N., & Fox, E. (2000). The role of perceptual load in negative priming. *Journal of Experimental Psychology: Human Perception and Performance, 26*(3), 1038.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General, 133*(3), 339.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Attention, Perception, & Psychophysics, 56*(2), 183-197.

Leibowitz, H., & Appelle, S. (1969). The effect of a central task on luminance thresholds for peripherally presented stimuli. *Human Factors, 11*(4), 387-391.

Lim, H.-M., Newstrom, B., Shankwitz, C., & Donath, M. (2000). *A Conformal Augmented Head Up Display for Driving Under Low Visibility Conditions.* Paper presented at the Proceedings of the 5th International Symposium on Advanced Vehicle Control, Ann Arbor, MI.

Lindgren, M., Stenberg, G., & Rosén, I. (1996). Effects of nicotine in visual attention tasks. *Human Psychopharmacology: Clinical and Experimental, 11*(1), 47-51.

Liu, Y.-C., & Wen, M.-H. (2004). Comparison of head-up display (HUD) vs. head-down display (HDD): driving performance of commercial vehicle operators in Taiwan. *International Journal of Human-Computer Studies, 61*(5), 679-697.

Livingston, M. A., Swan II, J. E., Gabbard, J. L., Höllerer, T. H., Hix, D., Julier, S. J., . . . Brown, D. (2003). *Resolving multiple occluded layers in augmented reality.* Paper presented at the Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality.

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. (2003). Visual perception of egocentric distance in real and virtual environments. *Virtual and adaptive environments, 11*, 21-46.

Macdonald, J. S., & Lavie, N. (2008). Load induced blindness. *Journal of Experimental Psychology: Human Perception and Performance, 34*(5), 1078.

Mack, A., & Rock, I. (1998). Inattentional blindness: Perception without attention. *Visual attention, 8*, 55-76.

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic science, 3*(1-12), 67-68.

Marrara, M. T., & Moore, C. M. (2000). Role of perceptual organization while attending in depth. Perception & Psychophysics, 62, 786–799.

Maylor, E. A., & Lavie, N. (1998). The influence of perceptual load on age differences in selective attention. *Psychology and aging, 13*(4), 563.

McIlwain, J. T. (1964). Receptive fields of optic tract axons and lateral geniculate cells: peripheral extent and barbiturate sensitivity. *Journal of Neurophysiology*, *27*(6), 1154-1173.

McKee, S. P., & Taylor, D. G. (2010). The precision of binocular and monocular depth judgments in natural settings. *Journal of vision, 10*(10), 5-5.

Millodot, M. (1972). Variation of visual acuity in the central region of the retina. *The British journal of physiological optics, 27*(1), 24.

Miura, T. (1990). Active function of eye movement and useful field of view in a realistic setting.

Neisser, U., & Becklen, R. (1975). Selective looking: Attending to visually specified events. *Cognitive psychology, 7*(4), 480-494.

Neveu, M. M., & Jeffery, G. (2007). Chiasm formation in man is fundamentally different from that in the mouse. *Eye, 21*(10), 1264-1270. doi:10.1038/sj.eye.6702839

Norman, D. A. (1968). Toward a theory of memory and attention. *Psychological review, 75*(6), 522.

Nuthmann, A. (2014). How do the regions of the visual field contribute to object search in real-world scenes? Evidence from eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 40*(1), 342.

O'Shea, R. P., Blackburn, S. G., & Ono, H. (1994). Contrast as a depth cue. *Vision Research, 34*(12), 1595-1604.

Okabayashi, S., Furukawa, M., Sakata, M., & Hatada, T. (1992). Driver's Ability to Recognize the Forward View and Head-Up Display Images in Automobiles. *Journal of Light & Visual Environment, 16*(2), 2_13-12_21.

Osterberg, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta ophthalmol, 13*, 6-97.

Owsley, C., Ball, K., & Keeton, D. M. (1995). Relationship between visual sensitivity and target localization in older adults. *Vision Research*, *35*(4), 579-587.

Owsley, C., Ball, K., McGwin Jr, G., Sloane, M. E., Roenker, D. L., White, M. F., & Overley, E. T. (1998). Visual processing impairment and risk of motor vehicle crash among older adults. *Jama*, *279*(14), 1083-1088.

Owsley, C., Ball, K., Sloane, M. E., Roenker, D. L., & Bruni, J. R. (1991). Visual/cognitive correlates of vehicle accidents in older drivers. *Psychology and aging, 6*(3), 403.

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.

Pashler, H. E., & Sutherland, S. (1998). *The psychology of attention* (Vol. 15): MIT press Cambridge, MA.

Pauzie, A. (2015). *Head Up Display in automotive: a new reality for the driver.* Paper presented at the International Conference of Design, User Experience, and Usability.

Pavel, M., Wang, G., & Li, K. (2003). *Augmented cognition: Allocation of attention.* Paper presented at the System Sciences, 2003. Proceedings of the 36th Annual Hawaii International Conference on.

Perry, V. H., & Cowey, A. (1985). The ganglion cell and cone distributions in the monkey's retina: implications for central magnification factors. *Vision Research, 25*(12), 1795-1810.

Petzoldt, T. (2011). Theoretical and Methodological Issues in Driver Distraction.

Philbeck, J. W., & Loomis, J. M. (1997). Comparison of two indicators of perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance, 23*(1), 72.

Pisano, E. D., Fuchs, H., State, A., Livingston, M. A., Hirota, G., Garrett, W. F., & Whitton, M. C. (1998). Augmented reality applied to ultrasound-guided breast cyst aspiration. *Breast Dis, 10*(3-4), 221-230.

Pitts, M. J., Hasedži´c, E., Skrypchuk, L., Attridge, A., and Williams, M. (2015) Adding depth: Establishing 3d display fundamentals for automotive applications. *Tech. Rep. 2015-01-0147, SAE International*,

Plavšic, M., Duschl, M., Tönnis, M., Bubb, H., & Klinker, G. (2009). Ergonomic design and evaluation of augmented reality based cautionary warnings for driving assistance in urban environments. *Proceedings of Intl. Ergonomics Assoc*.

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.

Poitschke, T., Ablaßmeier, M., Rigoll, G., Bardins, S., Kohlbecher, S., & Schneider, E. (2008). *Contact-analog information representation in an automotive head-up display.* Paper presented at the Proceedings of the 2008 symposium on Eye tracking research & applications.

Polyak, S. L. (1941). *The retina*: JSTOR.

Posner, M. I. (1980). Orienting of attention. *Quarterly journal of experimental psychology, 32*(1), 3-25.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual review of neuroscience, 13*(1), 25-42.

Posner, M. I., & Rothbart, M. K. (2007). Research on attention networks as a model for the integration of psychological science. *Annu. Rev. Psychol., 58*, 1-23.

Previc, F. H. (1990). Functional specialization in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences, 13*(3), 519-542.

Rafal, R., Gershberg, F., Egly, R., Ivry, R., Kingstone, A., & Ro, T. (1996). Response channel activation and the lateral prefrontal cortex. *Neuropsychologia, 34*(12), 1197-1202.

Rees, G., Frith, C. D., & Lavie, N. (1997). Modulating irrelevant motion perception by varying attentional load in an unrelated task. *Science, 278*(5343), 1616-1619.

Reifinger, S., Wallhoff, F., Ablassmeier, M., Poitschke, T., & Rigoll, G. (2007). Static and dynamic hand-gesture recognition for augmented reality applications. *Human-Computer Interaction. HCI Intelligent Multimodal Interaction Environments*, 728-737.

Reingold, E. M., & Stampe, D. M. (1999). Saccadic inhibition in complex visual tasks. In *Current oculomotor research* (pp. 249-255): Springer.

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.

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.

Rijsdijk, J., Kroon, J., & Van der Wildt, G. (1980). Contrast sensitivity as a function of position on the retina. *Vision Research, 20*(3), 235-241.

Ringer, R. V., Throneburg, Z., Johnson, A. P., Kramer, A. F., & Loschky, L. C. (2016). Impairing the useful field of view in natural scenes: Tunnel vision versus general interference. *Journal of vision, 16*(2), 7-7.

Roberts, K. L., Allen, H. A., Dent, K., & Humphreys, G. W. (2015). Visual search in depth: The neural correlates of segmenting a display into relevant and irrelevant three-dimensional regions. *NeuroImage, 122*, 298 305

Rock, I., & Gutman, D. (1981). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7*(2), 275.

Roda, C., & Thomas, J. (2006). Attention aware systems: Theories, applications, and research agenda. *Computers in Human Behavior, 22*(4), 557-587.

Rolland, J. P., Gibson, W., & Ariely, D. (1995). Towards quantifying depth and size perception in virtual environments. *Presence: Teleoperators & Virtual Environments, 4*(1), 24-49.

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: Teleoperators and Virtual Environments, 11*(6), 610-625.

Rovamo, J., Virsu, V., & Näsänen, R. (1978). Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature, 271*(5640), 54-56.

Sanders, A. F. (1970). Some aspects of the selective process in the functional visual field. *Ergonomics, 13*(1), 101-117.

Sands, J., & Lawson, S. W. (2003). *Towards a usable stereoscopic augmented reality interface for the manipulation of virtual cursors.* Paper presented at the Proceedings of the 2nd IEEE/ACM International Symposium on Mixed and Augmented Reality.

Sands, J., Lawson, S. W., & Benyon, D. (2004). *Do we need stereoscopic displays for 3D augmented reality target selection tasks?* Paper presented at the Information Visualisation, 2004. IV 2004. Proceedings. Eighth International Conference on.

Saunders, J. A., & Knill, D. C. (2004). Visual feedback control of hand movements. *Journal of Neuroscience, 24*(13), 3223-3234.

Schieber, F., & Benedetto, J. M. (1998). *Age differences in the functional field-of-view while driving: A preliminary simulator-based study.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Seiple, W., Szlyk, J. P., Yang, S., & Holopigian, K. (1996). Age-related functional field losses are not eccentricity dependent. *Vision Research, 36*(12), 1859-1866.

Seya, Y., Nakayasu, H., & Yagi, T. (2013). Useful field of view in simulated driving: Reaction times and eye movements of drivers. *i-Perception, 4*(4), 285-298.

Sharman, R. J. (2014). *Cue combination of colour and luminance in edge detection.* University of Nottingham,

Singh, G., Swan II, J. E., Jones, J. A., & Ellis, S. R. (2010). *Depth judgment measures and occluding surfaces in near-field augmented reality.* Paper presented at the Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization.

Smith, A., & Edgar, G. (1990). The influence of spatial frequency on perceived temporal frequency and perceived speed. *Vision Research, 30*(10), 1467-1474.

Sojourner, R. J., & Antin, J. F. (1990). The effects of a simulated head-up display speedometer on perceptual task performance. *Human Factors, 32*(3), 329-339.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological monographs: General and applied, 74*(11), 1.

Steinfeld, A. (1995). Driver response times to full-windshield, head-up displays for navigation and vision enhancement.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of vision, 11*(5), 13-13.

Surdick, R. T., Davis, E. T., King, R. A., Corso, G. M., Shapiro, A., Hodges, L., & Elliot, K. (1994). *Relevant cues for the visual perception of depth: is where you see it where it is?* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

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 and Virtual Environments, 6*(5), 513-531.

Swan, J., 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.* Paper presented at the Virtual Reality Conference, 2006.

Swan, J. E., Jones, A., Kolstad, E., Livingston, M. A., & Smallman, H. S. (2007). Egocentric depth judgments in optical, see-through augmented reality. *IEEE transactions on visualization and computer graphics, 13*(3), 429-442.

Teichner, W. H., Kobrick, J. L., & Dusek, E. R. (1956). Effects of target separation and distance on commonplace binocular depth discrimination. *JOSA, 46*(2), 122-125.

Teichner, W. H., Kobrick, J. L., & Wehrkamp, R. F. (1955). The effects of terrain and observation distance on relative depth discrimination. *The American journal of psychology, 68*(2), 193-208.

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 and Virtual Environments, 13*(5), 560-571.

Thorpe, S., Gegenfurtner, K., Fabre-Thorpe, M., & Bülthoff, H. (1999). *Categorisation of complex natural images in extreme peripheral vision.* Paper presented at the 22nd European Conference on Visual Perception.

Todd, J. T., Thaler, L., & Dijkstra, T. M. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research, 45*(12), 1501-1517.

Treisman, A., & Geffen, G. (1967). Selective attention: Perception or response? *The Quarterly journal of experimental psychology, 19*(1), 1-17.

Treisman, A. M. (1969). Strategies and models of selective attention. *Psychological review, 76*(3), 282.

Tufano, D. R. (1997). Automotive HUDs: The overlooked safety issues. *Human Factors, 39*(2), 303-311.

Tun, P. A., Wingfield, A., & Lindfield, K. C. (1997). Motor-speed baseline for the digit-symbol substitution test. *Clinical Gerontologist: The Journal of Aging and Mental Health*.

Turvey, M., & Kravetz, S. (1970). Retrieval from iconic memory with shape as the selection criterion. *Perception & Psychophysics, 8*(3), 171-172.

Underwood, G., Crundall, D., & Chapman, P. (2002). Selective searching while driving: the role of experience in hazard detection and general surveillance. *Ergonomics, 45*(1), 1-12.

Van de Weijgert, E. (2003). Foveal load and peripheral task performance: tunnel vision or general interference. *Visual search, 2*, 341-348.

Virsu, V., & Rovamo, J. (1979). Visual resolution, contrast sensitivity, and the cortical magnification factor. *Experimental Brain Research, 37*(3), 475-494.

Wallach, H., Gillam, B., & Cardillo, L. (1979). Some consequences of stereoscopic depth constancy. *Perception & Psychophysics, 26*(3), 235-240.

Walraven, J., & Janzen, P. (1993). TNO stereopsis test as an aid to the prevention of amblyopia. *Ophthalmic and Physiological Optics, 13*(4), 350-356.

Watson, J. D., Myers, R., Frackowiak, R. S., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., . . . Zeki, S. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral cortex, 3*(2), 79-94.

Wechsler, D. (1981). *WAIS-R manual: Wechsler adult intelligence scale-revised*: Psychological Corporation.

Weinberg, G., Harsham, B., & Medenica, Z. (2011). *Evaluating the usability of a head-up display for selection from choice lists in cars.* Paper presented at the Proceedings of the 3rd International Conference on Automotive User Interfaces and Interactive Vehicular Applications.

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica, 41*(1), 67-85.

Wickens, C. D., & Alexander, A. L. (2009). Attentional tunneling and task management in synthetic vision displays. *The International Journal of Aviation Psychology, 19*(2), 182-199.

Wickens, C. D., Gordon, S. E., Liu, Y., & Lee, J. (1998). An introduction to human factors engineering.

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.* Paper presented at the Proceedings of the 1st Symposium on Applied perception in graphics and visualization.

Willemsen, P., & Gooch, A. A. (2002). *Perceived egocentric distances in real, image-based, and traditional virtual environments.* Paper presented at the Virtual Reality, 2002. Proceedings. IEEE.

Williams, L. J. (1988). Tunnel vision or general interference? Cognitive load and attentional bias are both important. *The American journal of psychology*, 171-191.

Williams, L. J. (1989). Foveal load affects the functional field of view. *Human Performance, 2*(1), 1-28.

Wither, J., & Hollerer, T. (2005). *Pictorial depth cues for outdoor augmented reality.* Paper presented at the Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on.

Witmer, B. G., & Kline, P. B. (1998). Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments, 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.

Wolfe, J. M. (1999). Inattentional amnesia. *Fleeting memories, 17*, 5.

Wolfe, J. M., & Cave, K. R. (1999). The psychophysical evidence for a binding problem in human vision. *Neuron, 24*(1), 11-17.

Wong, B. P., Woods, R. L., & Peli, E. (2002). Stereoacuity at distance and near. *Optometry & Vision Science, 79*(12), 771-778.

Wright, J. v. (1968). Selection in visual immediate memory. *The Quarterly journal of experimental psychology, 20*(1), 62-68.

Yang, Z., & Purves, D. (2003). A statistical explanation of visual space. *Nature neuroscience, 6*(6), 632-640.

Yuille, A. L., & Grzywacz, N. M. (1989). A winner-take-all mechanism based on presynaptic inhibition feedback. *Neural Computation, 1*(3), 334-347.