Phenomenal regression as a potential metric of veridical perception in virtual environments

Kevin William Elner MSc
Department of Computer Science
University of Hull

This thesis is submitted for the degree of
Doctor of Philosophy

2015

For my beautiful boy
Malcolm

Acknowledgements

I would like to acknowledge...

The University of Hull Computer Science department for the doctoral student scholarship and opportunity to pursue the PhD.

My supervisor Dr Helen Wright who patiently mentored me through the PhD process, guiding me to conduct robust scientific research.

My panel chair Derek Wills and technical advisor James Ward for their encouragement and enthusiasm for the research topic.

The departmental systems team — Mark Bell, Andrew Hancock and Adam Hird. A special thanks to Mike Bielby for providing the technical support that enabled my research.

All the members of the department who kindly participated in the experiments.

My darling wife Susan for her unconditional support and encouragement.

Abstract

It is known that limitations of the visual presentation and sense of presence in a virtual environment (VE) can result in deficits of spatial perception such as the documented depth compression phenomena. Investigating size and distance percepts in a VE is an active area of research, where different groups have measured the deficit by employing skill-based tasks such as walking, throwing or simply judging sizes and distances. A psychological trait called phenomenal regression (PR), first identified in the 1930s by Thouless, offers a measure that does not rely on either judgement or skill. PR describes a systematic error made by subjects when asked to match the perspective projections of two stimuli displayed at different distances. Thouless' work found that this error is not mediated by a subject's prior knowledge of its existence, nor can it be consciously manipulated, since it measures an individual's innate reaction to visual stimuli. Furthermore he demonstrated that, in the real world, PR is affected by the depth cues available for viewing a scene. When applied in a VE, PR therefore potentially offers a direct measure of perceptual veracity that is independent of participants' skill in judging size or distance. Experimental work has been conducted and a statistically significant correlation of individuals' measured PR values (their 'Thouless ratio', or TR) between virtual and physical stimuli was found. A further experiment manipulated focal depth to mitigate the mismatch that occurs between accommodation and vergence cues in a VE. The resulting statistically significant effect on TR demonstrates that it is sensitive to changes in viewing conditions in a VE. Both experiments demonstrate key properties of PR that contribute to establishing it as a robust indicator of VE quality. The first property is that TR exhibits temporal stability during the period of testing and the second is that it differs between individuals. This is advantageous as it yields empirical values that can be investigated using regression analysis. This work contributes to VE domains in which it is desirable to replicate an accurate perception of space, such as training and telepresence, where PR would be a useful tool for comparing subjective experience between a VE and the real world, or between different VEs.

Publication

Elner, K. and Wright, H. (2015). Phenomenal regression to the real object in physical and virtual worlds. *Virtual Reality*, 19(1):21–31.

Contents

List of Figures

List of Tables

Terminology

Acronyms / Abbreviations

AR	Augmented Reality
AV	Augmented Virtuality
FOV	Field of View
HIVE	Hull Immersive Virtual Environment
HMD	Head Mounted Display
IPD	Inter-Pupillary Distance
LCD	Liquid Crystal Display
LED	Light Emitting Diode
LSID	Large Screen Immersive Display
MR	Mixed Reality
PR	Phenomenal Regression
TR	Thouless Ratio
VE	Virtual Environment

Nomenclature

P	Phenomenal character of perception
R	Real physical character of an object
S	Stimulus character cast upon the retina of the eye

Glossary

Augmented Reality	Thomas (2012) defines augmented reality (AR) as "*the registration of computer-generated graphical information over a user's view of the physical world*"
Augmented Virtuality	The majority of the environment is virtual, but retains some real objects. The virtuality of the virtual scene is augmented by the presence of these real-world objects
Character	An attribute or property of an object e.g. the characteristics of size or shape of an object
Dioptre	A unit to measure the refractive power of a lens. Lens power is equal to the reciprocal of the focal length and, therefore, calculated by $1/f$ where f is the focal length in meters
Interpupillary Distance	The distance of separation between an individuals eyes that is measured from pupil to pupil
Mixed Reality	Synthetic spaces are not necessarily fully virtual as they merge objects from both the physical and virtual realities within a single shared space. Milgram and Kishino (1994) refer to these shared spaces as mixed reality and describe them as a subset of virtual reality that sits along a virtuality continuum, the extremities being VE and AR
Objective size	Objective size is actual measurable size of an object i.e. while a coin held at arms-length and the moon may subtend at the same visual angle to the eye we know the objective size of the moon is far larger
Perspective size	Perspective size is the the visual angle subtended to the eye i.e. while a coin held at arms-length is physically much smaller than the moon both may subtend at the same visual angle to the eye and perspectively look the same size
Phemonenal Regression	Thouless (1931a) found that participants make a systematic error when asked to match the perceived characteristics of two

stimuli. It was found that the perceived (phenomenal *P*) characters of objects in the world do not match those of retinal stimulation (stimulus *S*), but are influenced by the character of the real (*R*) physical attributes of the object. Thouless found that *P* shows tendency from *S* towards *R* and referred to this as phenomenal regression to the real object

Reliability	A good metric would need to be reliable to produce consistent results across multiple measures when there is no evidence of change in the testing conditions — participant and environment. A reliable metric also should have good temporal stability, that is, be stable for the period between testing two comparators i.e. a VE and physical space
Responsiveness	To be truly useful any metric should be responsive (sensitive) enough to detect multiple levels of change
Thouless Ratio	A normalised measure of phenomenal regression. A TR value of zero means no regression whereas at unity regression is complete
Validity	The obvious requirement of any metric would be that it is valid, meaning that it actually does measure what it's supposed to
Virtual Environment	Ellis (1994) provides a useful definition of a virtual environment (VE) as "*[an] interactive, virtual image display enhanced by special processing and by non-visual display modalities, such as auditory and haptic [touch], to convince users they are immersed in a synthetic space*"

Chapter 1

Introduction

This thesis documents the research conducted to establish a psychological trait called phenomenal regression as a potential metric of veridical perception in virtual environments (VEs). Such a metric would be a useful yardstick to measure the deficit in visual perception in a VE and as a comparator between different VE configurations.

The goal of this chapter is to introduce the reader to the main topics of the research and to establish the motivation for conducting it. To that end, section 1.1 highlights why VEs are relevant today and warrant research attention. The section gives examples of current application domains, discussing the unique capabilities of VEs along with advantages of adopting VE technologies and finally the recent events that have stimulated demand and growth.

Section 1.2 follows by drawing attention to the need for a broad measure of visual perception quality due to the limitations of a VE in replicating the sensory stimulation of the real world. These limitations cause deficits in visual perception of space, which in turn affects a user's innate spatial skill in judging shape, size and distance. The section outlines the desired traits of a visual perception metric for VEs, contrasted against existing approaches for measuring user experience in a VE.

Section 1.3 states the aim of the research and lists the associated objectives to prove it.

The final section 1.4 provides the research plan and a roadmap of the remaining chapters of the thesis outlining the progress of research.

1.1 Relevance of VEs

This section makes the case that VEs are a relevant technology that are worthy of research attention, with a discussion of the current applications of VE, benefits of VE and the recent factors that suggest future growth and adoption of VE technology.

1.1.1 Current applications

VE technology has been adopted within numerous sectors of industry, commerce and academia. The applications of VE can be broadly categorised, based on function, into visualisation, training and education, manufacturing, design, therapy and tele-presence.

Using VEs for visualisation of data offers improved interaction and exploration of the data in three dimensions. Data visualisation in VEs has been particularly useful and widely adopted for geographical data as by Harding (2004), meteorological data as by Ziegeler et al. (2001) and mathematical data as by Klimenko et al. (2002). It has also been successfully used in genomics (Adams et al., 2002), molecular biology (Chastine et al., 2005), haematology (Pivkin et al., 2004) and neurological MRI (Lo et al., 2007; Zhang et al., 2001). It is also common to adopt the use of VEs for cultural heritage projects that involve the reconstruction of historical sites (Yue et al., 2006).

The VE is a firmly established component of modern training systems within numerous organisational domains such as aerospace, medicine and the military. The training can be focused on a particular task such as flight control (QinetiQPlc, 2013), medical surgery (Aggarwal et al., 2007) and firearm shooting (McCullum, 2011) or it can be situational, such an aircraft pilot practising landing protocol and communicating with air traffic control (VirtualAviationLtd., 2013). Training using VE is widely used within the process industries to train people for upstream oil and gas roles, for example off-shore drilling (Maersk, 2014). The United States Department of Energy, in the public sector, uses VE to train power plant operators on gasification processes (Mainwaring, 2012). VEs are also used for collaborative training such as simulation of military command and control exercises and have begun to penetrate the commercial consumer market as well. For example, the BSM driving school offers its students the option to train in a virtual car simulator before driving on the road (BSM, 2008). VirtualAerospace (2014) have even made their professional VE flight training simulators available for hire to the general public for recreational purposes.

Yet another industry that has embraced VE, specifically augmented reality (AR) technology to gain a competitive edge, is the retail industry. For instance, IKEA (2013, 2014) has capitalised on the ubiquity of smart-phones and tablet PC devices, offering its customers an AR application that enables viewing and arranging virtual furniture in their home. Similarly, clothing brands are starting to deploy AR fitting in stores, using the analogy of looking in a mirror, where customers can view themselves wearing virtual clothes (Yuan et al., 2013).

We can speculate that the advantage of such applications increase the conversion rate from browsing to a sale.

VEs are also used to provide an emerging form of therapy for pre-emptive treatment of post traumatic-stress disorder (PTSD) commonly caused by intense military activity (Rizzo et al., 2012). Hoffman et al. (2011) has found that immersing burns patients in a VE can be just as effective at controlling pain as traditional drugs. Other examples of VEs used for treatment include the treatment of phobias like arachnophobia (Juan et al., 2005), acrophobia (Lear, 1997) as well as social phobias like public speaking (Slater et al., 1999).

1.1.2 Benefits of adoption

A VE immerses a user in a computer-generated simulation of reality. The primary advantage of a VE lies in its ability to replicate the experience of the physical world. However, a VE need not be restricted to the constraints of the physical world or the user's physical ability. A VE could allow for users to interact with an environment in a manner that otherwise would be impossible. For instance a VE can allow for users to interact with an environment using hand gestures or navigate through it using only thoughts (Friedman et al., 2007).

VE has been widely adopted for modern training systems as it can offer a more cost effective solution when compared to traditional training approaches providing savings on consumables, waste and the wear and tear on equipment. A further advantage is that VE training systems are configurable to very specific learning and training requirements with modifiable difficulty and can record all actions for improved skills analysis and performance evaluation. Additionally, VEs allow for the safe rehearsal of scenarios that have significant consequences, are non-repeatable, are difficult to recreate and/or rarely occur in the real world. Another advantage is that VE enables new training ability by visualising normally unseen artefacts such as radiation (Vertual, 2014).

VE even allows the expertise of a user to be transported from one geographical location to another with the use of robotics to interact with the environments, hence called 'tele-presence'. This is particularly useful to create safer working in hazardous environments like nuclear reactors. Tele-presence tends to rely on a 'live' video feed from the remote environment, while VE adds the additional ability of a digitally reconstructed environment from sensors like sonar devices. Applications of tele-presence such as tele-medicine rely on haptic interfaces to enable surgeons to conduct surgical procedures remotely.

Tele-medicine, a particular domain / application of tele-presence that involves transporting medical expertise, is an active area of research with these kinds of applications relying on haptic feedback.

1.1.3 Recent evidence of demand for VEs

Historically, the use of VE hardware for consumer applications such as gaming has been a costly option with a lack of third party developer support. Since 2013, OculusVR are attempting to address this by getting affordable VE technology into the hands of gamers as well as supporting game developers to make use of the technology. With 9,522 backers the project raised $2,437,429 using the Kickstarter crowd funding platform (Kickstarter, 2014; OculusVR, 2012, 2014). Many of the current best selling computer games are already designed to be experienced from a first-person perspective so it is a natural progression to adopt VE hardware as demonstrated, for example, by the Battlefield 3 simulator (TheGadgetShow, 2011). Gamesutra (2013) reported that worldwide revenues of the games market grew to $70.4 billion in 2013, a growth of 6%. This huge market for VE technologies in a growing games market should result in a reduction of hardware costs due to the economies of scale and in turn provide some stimulus to growth of the technology. The latest development of Facebook acquiring OculusVR (Lee, 2014) has further implications on the widespread use of VE technology in the non-gaming social media arena.

As per the predictions made by the ‘Global Military Simulations and Virtual Training Market 2011–2021’ report published by ICD Research (Field, 2012), defence budgets appear to be negatively affected by the global financial crisis of 2007–08. The cost savings offered by military simulators seem to be driving the demand for VE as governments seek cost-effective alternatives to traditional training.

A recent event that has resulted in the increased use of VE for training in the oil industry is evident from the following statement by Mainwaring (2012) *“Shell, mindful of the huge cost [financial and environmental] and damage to the reputation of rival oil firm BP due to the Macondo disaster adopted real-time training methods that employ aspects of virtual reality”*.

The recent push towards using VE technology in computer gaming and recreation domains, as well as recent socio-economic/political factors influencing the increased demand in the training domain, are both likely to result in the growth of the market for the use of VE technology.

Given the wide range of domains, the benefits of adoption and the recent developments that have led to adoption, it is a fair assumption that there will be continued growth in the adoption of VEs for the foreseeable future.

1.2 Metric of VE quality

This section begins by establishing the need for a measure of the quality of visual perception in a VE. This is followed by a short review the existing qualitative and quantitative approaches for measuring user experience in a VE. The desired traits of a metric for veridical perception in a VE are outlined, contrasted against the existing VE measures.

1.2.1 Establishing the need for a VE quality metric of visual perception

A VE can generate a synthetic space that is visually perceivable by its users by replicating the visual cues normally experienced in the real world. However, the accuracy of this simulated spatial experience is limited by its inability to stimulate the human visual senses to the same extent as the physical world. The resulting difference in stimulation modifies the user's innate physiological and psychological responses leading to a distorted perception of space. Such deficits in spatial perception can have an effect on innate spatial abilities such as judgement of size, distance and shape. As such, a VE does not accurately replicate a genuine experience of physical space as discussed in further detail in Chapter 2.

Accurate spatial perception may not be necessary for many of the previously mentioned VE applications. However, it may have a significant impact upon the effectiveness of VE training applications. This is especially important for those applications which aim to improve the user's ability to perform tasks that rely on innate spatial skills.

The effectiveness of a training task is assessed in terms of 'transfer of training' or simply 'transfer'. Pennington et al. (1995) defines transfer as is the extent of retention and application of knowledge, skills and attitudes from the training environment to the target environment. It has been theorised that a factor that may affect transfer is the principle of identical elements (Baldwin and Ford, 1988). This concept that was first conceived by Thorndike and Woodworth (1901); it predicts that transfer will be maximised when there is identical stimulus and response elements.

Although there has been some interest in the transfer of VE training since the advent of VE training solutions, as stated by Bossard et al. (2008), transfer (in VE) is an elusive concept despite the research devoted to understanding it. Rose et al. (2000) highlighted that although there is some anecdotal evidence of transfer of VE training to the target task, there are relatively few empirical investigations. Similarly and more recently Muller et al. (2006) mention that modelling and simulation tools are often advertised as providing effective training when little actual assessment has been done.

Visual perception of a VE space is not identical to physical space and the resulting distortion could affect training effectiveness. This effect was demonstrated by the results of an experiment on training conducted by Judd (1908), where a group of children, ignorant of the principles of refraction, were asked to throw darts at a target submerged in 12 inches of water. At first the children struggled to perform the task, but over a short period they adapted and through a process of trial and error they developed a strategy to hit the target. The experiment was then repeated with the depth of the water reduced to 6 inches. The participants could not hit the target and were confused as to why they could not reuse their previous strategy. In other words, they were unable to adapt to the different spatial distortion caused by the shallower water. Similarly a VE user may adapt to overcome a deficit in spatial perception and learn to perform a task well. However, as with Judd's participants there may be difficulty transferring the learnt concepts/behaviour between different configurations of VE and the real world. This difficulty in transfer of training has an implication on VE training applications as the trainee expects and is expected to transfer the training and perform the task in a real world situation. This is an example where training is not effective and where it has the the opposite effect being detrimental.

There is a need for a measure of VE quality, as a yardstick to assess magnitude of perceptual deficits in a VE, allowing for comparisons between virtual and physical spaces and to mediate any deficits. VE training is a domain that would benefit from getting VE perceptually correct. A VE quality metric could potentially be used as one of the criteria for assessing the quality of VE training. However, it is important to note that the focus of this research is to measure the perceptual deficits in a VE and not to assess the effectiveness of VE training.

1.2.2 Existing approaches to measure VE experience

Before stating the desired traits of a VE quality metric it is necessary to review some of the existing approaches for measuring the quality of various aspects of user experience in a VE. These measures can be broadly categorised as being either qualitative or quantitative.

Qualitative research analyses descriptive data collected from the perspective of a participant or passive observer. Common approaches to collect qualitative data include questionnaire, interview, focus group discussion and observation. Qualitative data is subjective and, therefore, likely to be influenced by emotion or the state of mind of the informant. It is expressed in the informant's language and is interpreted by subjecting it to content and thematic analysis. Qualitative data provides nuanced and contextual results; for example human behaviour such as body language cannot simply be quantified with numbers, but must also be observed and expertly analysed. Qualitative data can also be analysed numerically by assigning it to an arbitrary scale.

Quantitative research collects numerical measurement data. It is uninfluenced by the informant and, therefore, objective. Quantitative data is analysed statistically to identify the effect of independent variables upon a dependant variable. Data is collected through controlled experiments that isolate and measure these variables.

The study of presence in VEs makes extensive use of qualitative techniques. Presence refers to experiencing the computer-generated environment rather than the actual physical locale (Witmer and Singer, 1998). Van Baren and IJsselsteijn (2004) provides a comprehensive review of methods adopted to measuring presence with the most commonly used appearing to be post-test questionnaires, of which many different variations have been devised. Typically these questionnaires collect ordinal data by asking participants to recall aspects of their experience of presence in a VE and rank it using a numeric scale. There are also continuous self-reporting measurement techniques whereby the participant, verbally or by means of some apparatus, continuously rate presence throughout the VE experience.

Quantitative physiological measures provide a means to study emotional responses in a VE. Examples are electrocardiogram (ECG) to record cardiovascular activity to study acrophobia (Juan and Pérez, 2009), galvanic skin response (GSR) for social presence studies (Slater et al., 2006) and facial electromyography (EMG) to measure effect of social agents (Philipp et al., 2012). Such measures, however, tend to rely on extreme circumstances, such as inducing anxiety or excitement, to obtain a measurable physiological response and it is not always possible to determine what is causing the effect. Electroencephalogram (EEG) have been used to measure patterns of brain activity to provide an objective quantitative comparison of the mental processing differences that exist between real and virtual images (Strickland and Chartier, 1997). EEG has also been used to measure presence (Azevedo, 2014).

Ruddle and Lessels (2006) behaviour study of way-finding in a VE used a combination of techniques. Performance and physical behaviour was measured using quantitative metrics; time taken, distance travelled, number of errors, head-movement and locomotion. Cognitive data was collected using qualitative techniques; questionnaire, verbal self-reporting and observation.

VE training applications measure the behaviour and ability of a trainee to perform a computer generated task. Examples in the medical domain include procedure training for paracentesis (Tzafestas et al., 2008) and suturing (O'Toole et al., 1999). Tool technique was measured by orientation and positioning, smoothness of motion and accuracy of skin puncture. Other measures include total time to complete the procedure, tissue damage and number of errors made.

Spatial perception studies in VE have used a verbal reporting measure of perceived distance estimates (Klein et al., 2009; Kunz et al., 2009; Piryankova et al., 2013), however, it can be difficult to for participants to accurately quantify this. As such, other studies have relied on measuring a visually guided action to quantify the distance estimate; walking (Kunz et al., 2009; Thompson et al., 2004; Willemsen et al., 2009; Willemsen and Gooch, 2002; Willemsen et al., 2008), throwing (Sahm et al., 2005) and reaching (Ellis and Menges, 1998; Singh et al., 2010). These measures rely on both the skill of making the visual judgement and performing the action. Kenyon et al. (2007) investigated size perception using a size comparison task. The participant is presented with a familiar object at a distance and asked adjusts its size until it appears correct. This removes the need for a visually guided action, but still relies on skill of judging the size of an object at distance.

1.2.3 Desired traits of a quality metric for veridical perception in a VE

It is desirable that a VE quality metric to measure veridical perception will be quantitative. Although subjective qualitative measures have their importance and place, they may not easily be reproduced. A objective quantitative measure that is uninfluenced by a participants current emotion state and factually verifiable is likely to be a good metric of perceptual deficit in VE and useful when comparing different VEs.

Unlike the previous measures of spatial perception a VE quality metric would ideally remove the need for the application of a judgement and motor-skill in the VE. An additional concern of using such skill-based measures is that any prevailing perceptual deficits that we wish to measure could arguably be compensated for as demonstrated in the real-world by Judd (1908).

Ideally then a metric that quantifies quality of spatial perception in a VE would be a direct measure (non-skill based) and have the basic properties of validity, reliability, responsiveness;

- **Validity** The obvious requirement of any metric would be that it is valid, meaning that it does actually measures what it's supposed to. This first requirement of validity of the metric is mentioned for the sake of completeness

- **Reliability** A good metric would need to be reliable to produce consistent results across multiple measures when there is no evidence of change in the testing conditions

— participant and environment. A reliable metric also should have good temporal stability, that is, be stable for the period between testing two comparators i.e. a VE and physical space

- **Responsiveness** To be truly useful any metric used as a basis of comparison between VE and physical spaces must be able to detect small changes. Therefore, the metric should be responsive (sensitive) enough to detect multiple levels of change

1.3 Research objectives

The general aim of the research is to identify a direct measure of VE quality that relies on participants' innate perceptual responses to the visual stimuli, and the following objectives have been identified to fulfil the research aim:

i. Identify a key element of spatial perception where there is a suspected deficit in the VE compared to the real world.

ii. Identify a directly measurable property capable of characterising subjective perceptual experience of visual stimuli.

iii. Within the scope of [i], investigate the performance of the property identified in [ii] to establish its value as a VE assessment metric.

1.4 Research plan and thesis roadmap

As stated above, the aim of the research is to identify a direct measure of VE quality that relies on participants' innate perceptual responses to the visual stimuli. The subject of Thouless work in 1932, before VEs, suggests such a metric is possible in a property called phenomenal regression (PR). PR was found to rely on innate perceptual responses to the visual stimuli. PR is attractive to this research since as it has been found to be sensitive to reduction in available visual depth cues in the real world. This research will test if PR is applicable to VE and if it is suitable for measuring the quality of a VE.

Chapter 2 presents the background literature review fulfilling research objective [i]. It covers the physiology and psychology of human visual perception, how VEs creates synthetic space, identified deficits of visual perception in VE and reviews empirical studies for measuring perception of size and distance in a VE. Visual display modalities and the different spaces VEs generate are classified to help establish the research scope.

Chapter 3 investigates the psychological trait phenomenal regression (PR) fulfilling research objective [ii]. It begins by describing PR, its numerical value the Thouless ratio (TR) and the methodology to measure it. The informed restating of PR is based upon both the original the work carried out by psychologist Robert Thouless and exploratory experimental work necessary to overcome barriers to understanding it. The chapter concludes by drawing together the valuable lessons for future experiments and discussing the presumptive properties of PR for this research.

Chapter 4 begins by re-stating the research aim and outlining how the work conducted thus far has furnished research objectives [i] and [ii]. Research objective [iii] is refined and the hypothesis of the thesis is stated. In order to fulfil objective [iii] and thus prove the hypothesis, experimental work must be conducted to the demonstrate key properties of TR that establish it as useful measure. These key properties are stated.

Chapter 5 presents a experiment conducted to compare TR between a virtual and physical stimulus. It includes the experiment design, methodology, results and a discussion. The results show a statistically significant correlation of TR between virtual and physical stimuli. It concludes by highlighting which key properties the result has furnished in order to prove the hypothesis. TR has been demonstrated to be a valid and reliable measure.

Chapter 6 presents an experiment conducted to assess the effect on TR of manipulating the accommodation-vergence mismatch. It includes the experiment design, methodology, results and a discussion. The result was statistically significant which indicates that TR is sensitive to changes of depth cues in the VE and, therefore, a responsive measure.

Chapter 7 concludes the thesis by summarising the research conducted. The demonstrated key properties of TR are stated as is the contribution this investigation brings to the VE field. The key properties of TR that have not been demonstrated are discussed along with the future directions of the research.

Chapter 2

Research background

This chapter presents the research background. Section 2.1 begins by summarising the physiology of sensation and the flow of visual data from the eye to the brain. After sensation comes the psychology of perception and a look at some of the different theories of how visual perception works. Finally the monocular and stereo visual cues of spatial perception are described.

Having described both the physiological and psychological aspects of visual spatial perception, section 2.2 broadly discusses how VE technologies artificiality recreate visual depth cues to generate visually perceivable spaces. It covers screen-parallax, perspective projection, viewing frusta, stereo-pair generation and systems for viewing stereo.

Section 2.3 highlights the problem that VE is not able to generate an accurate visual perception of space and its effect on spatial perception. It also describes the consequences of visual latency in VE, the conflict of accommodation-vergence cues and visual crosstalk. It finishes with a review of empirical studies of distance and size perception in VEs.

Section 2.4 then classifies the variety of spaces that a VE can generate by describing a virtuality continuum and the dimensions of transportation in a VE as well as discussing the various visual display modalities to clearly define the scope of the research.

Finally section 2.5 highlights how this chapter has fulfilled objective [i] of the research aim.

2.1 Overview of human visual perception

Perception is the acquisition and processing of sensory information in order to experience objects that make up the world. Spatial perception is therefore the awareness of an object and its relationship between both the perceiver and other objects. Spatial ability refers to the actual skill in perceiving the world — abilities such as tracking objects though space transformations, determining inter-object relationships and being able to mentally manipulate object orientations in that space.

The following subsections give an overview of how the physiology and psychology of the human visual perceptual system works.

2.1.1 Physiology of the visual sensation — the eye and brain

To perceive we must also sense. Sensation is the ability of our sense organs to detect energy; the only visual sensor for humans are the eyes (Figure 2.1).

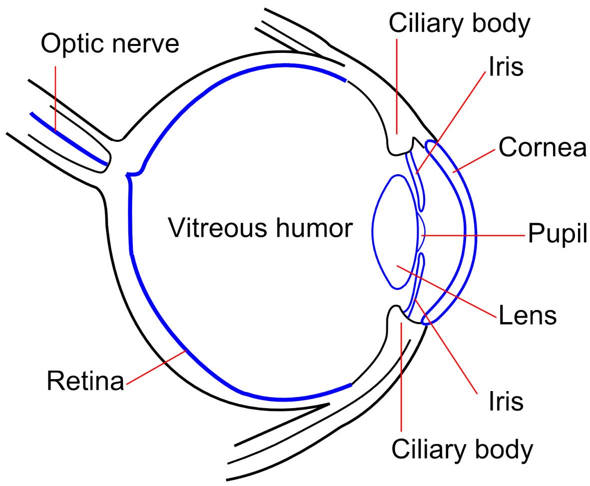
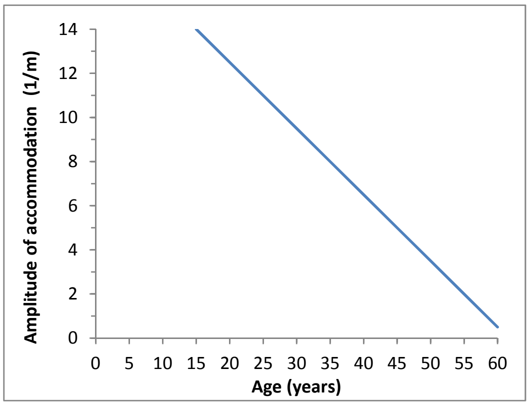


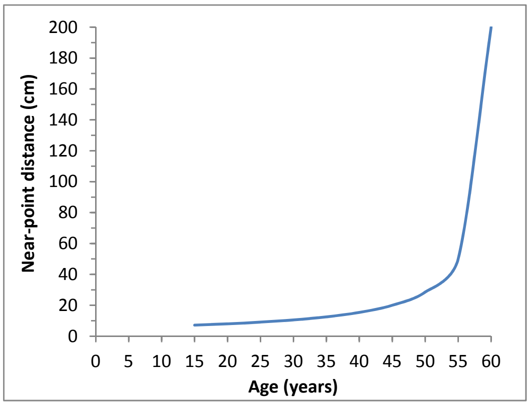
Fig. 2.1 Anatomy of the human eye

Objects in the world (stimuli) arouse the receptors of our eyes via light energy. Rays of light initially enter the outer eye through the cornea — the principal refractive element of the eye — and meet the iris. The iris acts as an aperture where light rays can only enter the inner eye via the pupil hole in the centre. It can dilate and shrink, varying the pupil size to regulate the amount of light that enters the eye, the depth of focus and quality of the retinal image. Rays of light enter the pupil and hit the crystalline lens.

The muscles of the ciliary body control the refractive properties of the lens allowing focusing — this action is termed accommodation or to accommodate. The change from relaxed to maximum ciliary exertion is called the amplitude of accommodation. As the eye ages there is a progressive decline in the amplitude of accommodation and, therefore, near-field vision (Figure 2.2). This is called Presbyopia and its cause has been attributed to atrophy of the ciliary muscles and the hardening of the crystalline lens with age (Charman, 2008).



(a)



(b)

Fig. 2.2 Graphs showing the effect of age on the accommodation of the eye. The values were calculated using the age-amplification formula by Hofstetter (1950). Graph (a) shows amplitude of accommodation plotted as a function of age. Between the ages of 15 to 60 the amplitude falls from 14 Dioptres to approximately zero, however, Mather (2009) says ability to accommodate is typically lost by 50 years. Graph (b) shows the near-point distance plotted as a function of age. The near-point distance of accommodation is the closest the eye can focus on when the focal power of the lens is maximised. After 40 years near-field vision quickly deteriorates and post 55 years it is lost.

Light is refracted by the lens onto the retina. The retina at the base of the eye samples the light admitted by the optical system of the eye (retinal image), processes it and transmits it to the brain via the optic nerve. Johansson (1975) says that the eye is basically an instrument for analysing the changes in light flux over time. The receptors in the retina do not capture light rather they measure changes in light. In other words the eye does not take a series of static images; it does not mimic a photography camera.

The optic nerve connects to the 'lateral geniculate nucleus' where the retinal data is mapped onto its surface. From there it is fed into the primary visual cortex of the brain. The primary flow of visual information from our senses is split into two streams of visual flow —- dorsal and ventral, see figure 2.3. The dorsal stream connects to a part of the brain which is believed to be responsible for analysing position, orientation and movement, while the ventral stream connects to a part of the brain that is believed to be for pattern discrimination.

2.1.2 Psychology of visual perception — theoretical models

Stimuli in the world arouse the receptors of our eyes and this information is transmitted to various areas of the brain that process it. How we actually form a visual perception is not

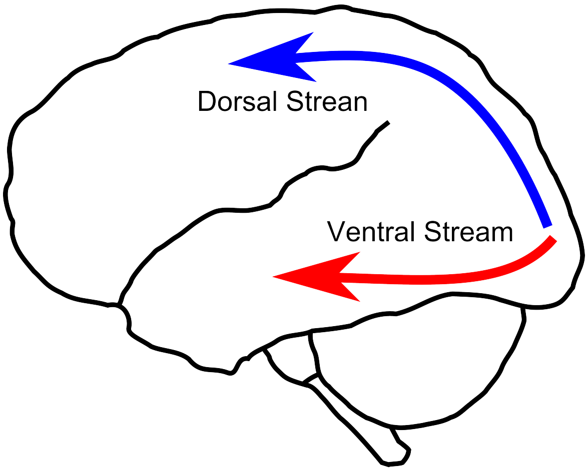


Fig. 2.3 Pathways of visual flow through the brain — dorsal and ventral

quite clear, but various theories have been proposed. Perceptual flow, the flow from sensation to perception, has been categorised to be either bottom-up or top-down. Bottom-up processing is based purely upon sensation and its goal is to form internal (self) representation of the world or objects. Top-down processing is purely cognitive and depends on knowledge stored in the brain to guide or 'make sense of' our visual senses. Gibson (1966) considered the senses as a complete perceptual system and that information is simply "picked up" and that the brain does not operate on the deliverances of sense or past experience. He rejects the requirement for any cognitive information and claims that all information necessary for visual perception is available in the environment. This has become known as an ecological approach to perception and flow is bottom-up. Gibson believed that the goal of perception leads to action without the need for a formal internal description of objects in the environment. Marr (1982) followed this ecological bottom-up approach, but felt that the goal of perception was in fact object recognition and not action. Braisby and Gellatly (2005) made an interesting synthesis, that both describe different aspects of ecological perception, unifying both ideas. The dorsal stream of the brain could be said to be concerned with 'where objects are?', which fits well with Gibson's perceptual goal of action. Similarly the ventral stream could be said to be concerned with 'what objects are?' which fits well with Marr's perceptual goal of object recognition.

Another theory of perception is by Gregory (1998). A top-down approach, it is considered a constructive approach to perception, illustrated in Figure 2.4. Gregory believes that sensory information is the basis of perception but that it is incomplete and that to be able to build or construct a complete picture, knowledge is necessary to interpret the sensory information. The incomplete sensory information forces our brain into making a series of perceptual hypotheses and that whichever is best supported by the sensory data is finally chosen.

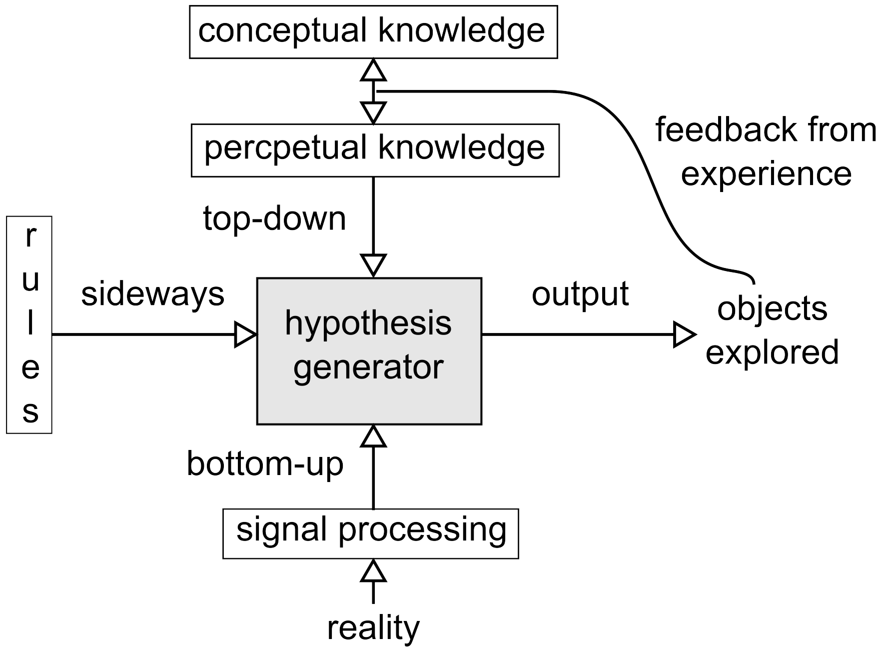


Fig. 2.4 Gregory's top down model of visual perception [Recreated from Gregory (1998)]

The criticism of Gregory's top-down approach is the question mark over how the hypothesis generator works. The sideways rules of Gregory's model refer to Gestalt. Gestalt thought is that objects are perceived as a whole according to the organisation of their parts or elements i.e. we recognise a person as a whole, not as a collection of individual pairs of arms and legs, a torso and head. Ellis (1991) supported the constructive approach to perception, that our sense of physical reality is largely a consequence of internal processing rather than developed from any immediate sensory information we receive. The sensory and cognitive interpretive systems are predisposed to processing incoming information in ways that normally result in correct interpretation of the external environment. The illusion of an enveloping environment depends on the extent to which all of these constructive processes are triggered. This suggests that the success of a VE to accurately generate a perceivable space depends on how well it can trigger the constructive process via the senses. Sensation is then a primary input of visual perception. Coren (2003) stated that our perception will always conform to our presumptions about the world and we will distort or warp our perception of reality to fit with those presumptions. Presumption comes from knowledge. This is illustrated by the Ames's room illusion (Figure 2.5) whereby to conform to our senses and belief that the room is square the perception of the size of the people in the room becomes distorted. This demonstrates how powerful knowledge is for visual perception.

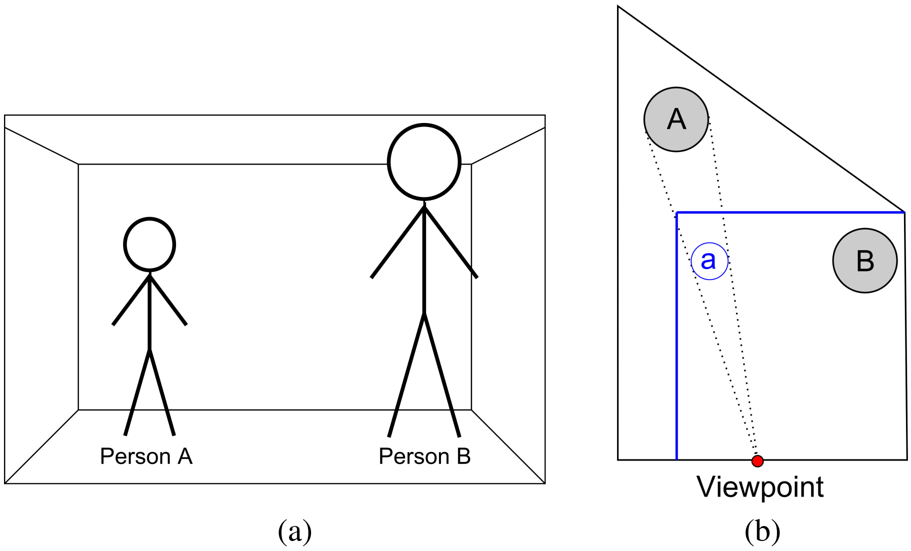


Fig. 2.5 The Ames room is viewed monocularly through pinhole at the viewpoint. The primary clues to the shape of the room are the perspective cues and a presumption of the rooms shape — knowledge (a) To a viewer of the Ames room, it appears rectangular in shape with person 'A' appearing smaller than person 'B' and the two individuals appear to be in line with each other. (b) A plan view of the Ames room shows the actual shape of the room and the actual positions of the two people. The blue line indicates the viewers perception of the rooms shape.

Synthesising both ecological approaches of Gibson and Marr and the constructive approach of Gregory, the process of visual perception can be simplified by the model shown in Figure 2.6. The inputs of this model are visual sensation and knowledge which are processed to form a visual perception. The output of model is a visual perception that will results in an action or recognition. Experience feeds back as knowledge. Within a VE the knowledge input can be considered to be the same as it would be in the real world, however, the sensation input is dependent on what stimulation the VE can provide. The quality of the output — action or object recognition — is therefore some function of the quality of the inputs — sensation and knowledge.

2.1.3 Visual cues of depth — monocular, stereo and ocular-motor

There are a number of monocular, stereo and ocular-motor cues that indicate depth.

Monocular cues like perspective indicate the distance of an object from the viewer. The objects in the world are projected onto the retina and objects that recede in distance from the viewer subtend a smaller retinal image on the eye occupying a progressively smaller

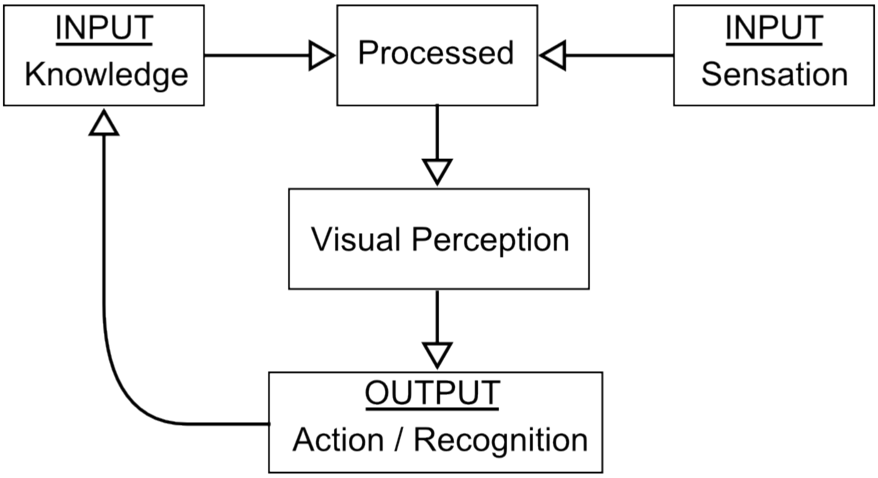


Fig. 2.6 A synthesis of the ecological and constructive models of visual perception, summarising the inputs and outputs.

portion of their field of view. Thus objects that are far away appear smaller than objects that are closer in distance and so parallel lines tend to converge as they recede into the distance (Figure 2.7).

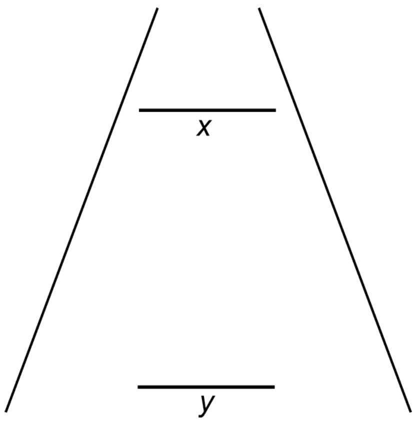


Fig. 2.7 This illustrates how receding parallel lines give clues of size. The lines x and y cast the same retinal stimulus upon our retina. However, the perspective cues make x appear further away than y and as a result x is perceived to be larger than y.

Other monocular cues of depth include pictorial cues like the gradient of textures, shading and shadows caused by lighting giving the impression of the position, orientation and the shape of objects (Figure 2.8). Thus more detail can be discerned on the surface of objects that are closer, while details of objects in the distance will appear less sharp .

Motion-parallax can be considered an additional pictorial cue of depth. When an observer moves, objects at varying distances from the observer will appear to move in differ-

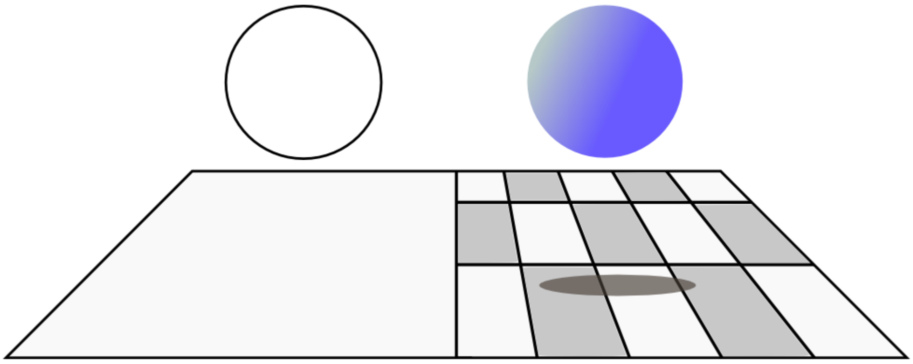


Fig. 2.8 Example of texture gradient, shading and shadow cues. The texture gradient give a clue to distance. The shading cue gives clue to shape. The shadow cue gives a clue to the objects position.

ent directions at different speeds (Coren et al., 2004). Objects in the foreground move faster than objects that are far away (Figure 2.9). This movement of objects serves as a cue to the relative distance of the objects, which is called motion-parallax.

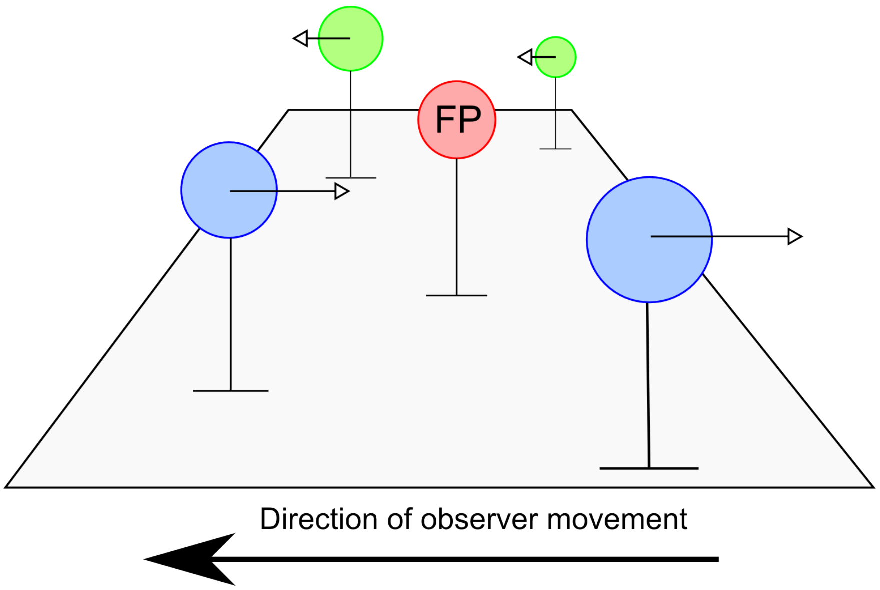


Fig. 2.9 Motion-parallax depth cue. When an observer fixates their view upon object FP (fixation point) and moves sideways to the left, in the observers field of view the foreground objects move to the right and the background object moves to the left. The direction and magnitude of movement is indicated by the arrows and its length.

Humans have two eyes that provide the brain with two offset sources of visual information — binocular disparity. This provides additional information cues to the brain's visual system that allow us to discern depth. Our perception of the world is a combination of the stimulation provided by both eyes, called stereo-vision, combining forms from the two sources to produce a single form that has three dimensions.

There are physiological depth cues that are a result of two (coupled) ocular motor cues accommodation and convergence — referred to as accommodation-vergence (Figure 2.10).

Accommodation is the movement of the ciliary muscles to focus the lens on the point of fixation. For a fixation point beyond 6m the lens is said to be focused at infinity as the lens is flat and, therefore, light rays entering the eye are parallel. According to Coren et al. (2004) beyond 3m the ciliary muscles are completely relaxed and thus accommodation ceases to provide any depth information. However, citations by Hershenson (1999) of Leibowitz's work suggest that it is only effective up to a distance of 1–2m. The focal range between those distances is 1/2 a Dioptre and beyond 2m the required change in focal strength diminishes (Figure 2.11 a) as does the usefulness of the accommodation cue.

Convergence is the movement of both eyes to fixate on a point in space. Palmer (2002) describes how at distances beyond 1.83–2.44m the angle changes very little as it approaches a limiting value of zero degrees (Figure 2.11 b). Similarly Mather (2009) says that by 2m 90% of the full range of vergence has been used up.

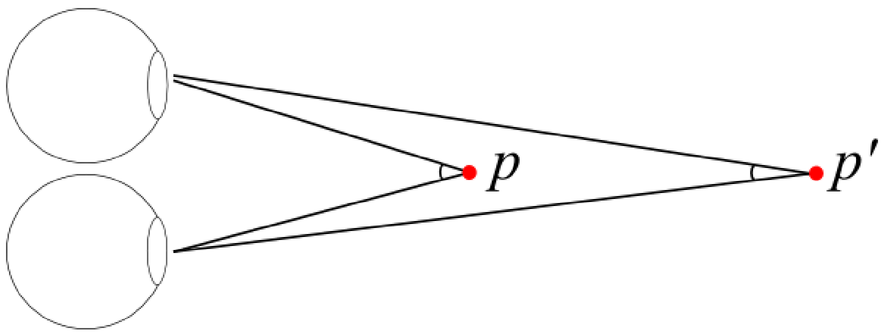
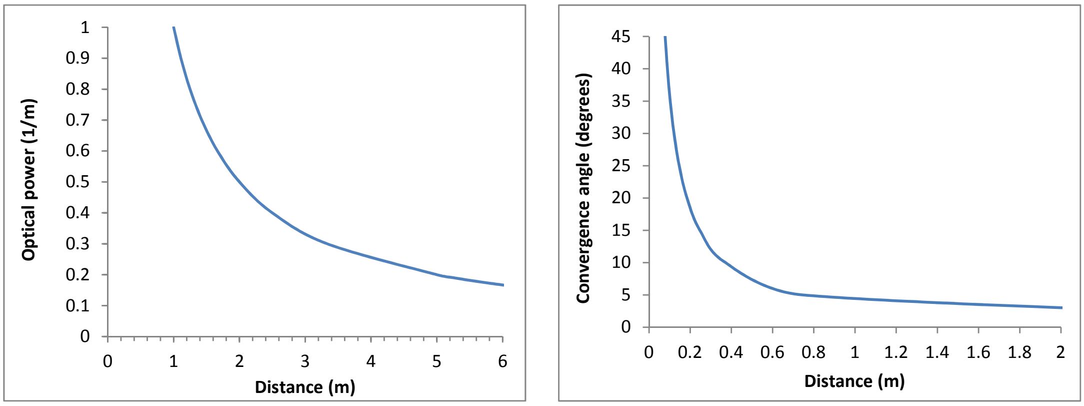


Fig. 2.10 Accommodation and convergence cues. When changing the fixation from p to p' the convergence angle of the eyes decreases and the lens of the the eye relaxes.

2.2 How VEs generate visually perceivable virtual spaces

VEs visually create virtual spaces by replicating the previous discussed monocular and stereo cues. The primary monocular cues are perspective and motion-parallax, but these are further enhanced by graphical techniques such as texturing, lighting and shading. Motion-parallax cues are made available by tracking the user's head position and orientation and re-calculating the perspective projection in real-time.



(a) (b)

Fig. 2.11 Graphs showing convergence and focal power of the eye as a function of distance. Graph (a) shows focal power as a function of distance. Graph (b) shows convergence angle as a function of distance.

The subsections that follow provide an brief overview of screen-parallax, perspective projection, calculating the frusta to generate a stereo-pair and display systems to view them.

2.2.1 Screen-parallax

Stereoscopy is achieved by rendering a left-eye and right-eye image — stereo-pair — of the 3D scene onto a projection plane. The separation of these images on the projection plane — screen-parallax — creates binocular disparity and, therefore, the illusion of depth. Positive-parallax describes the outward separation of the left-eye and right-eye images, objects appear to be at distances beyond the projection plane (Figure2.12 (a)). Zero-parallax is when there is no separation between the left-eye and righy-eye image, objects appear to be at the distance of the projection plane (Figure 2.12 (b)). Negative-parallax describes when the left-eye and right-eye images cross-inwards making objects to be in front of the projection plane (Figure 2.12 (c)).

2.2.2 Perspective projection and viewing frusta

To create a three dimensional view of a computer generated scene it is projected onto a two dimensional plane using a projection method. Using a perspective projection, of mapping three dimensional points on to a two dimensional pane, allows objects in the distance to appear smaller than objects close by similar to how a human eye views a scene. In computer graphics this is normally done by implementing a viewing frustum or clipping volume. The

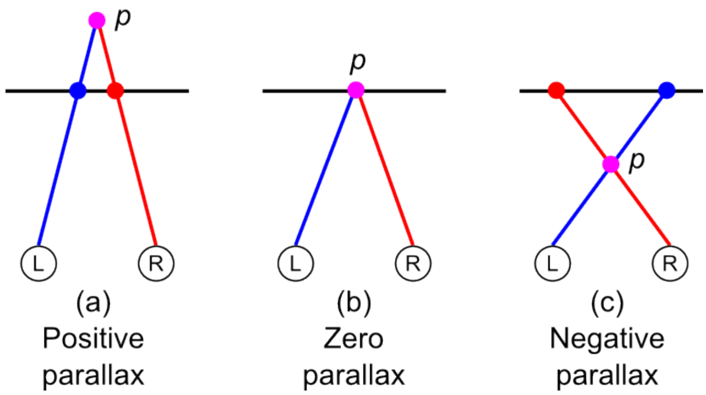


Fig. 2.12 Screen-parallax — how the separation of the left-eye (L) and right-eye (R) images on the projection plane (solid black line) changes the perceived (*p*) position of objects. (a) Positive-parallax (b) Zero-parallax (c) Negative-parallax.

shape ensures that objects in the distance fill a smaller portion of the field of view than objects that are close to the viewer (Figure 2.13)

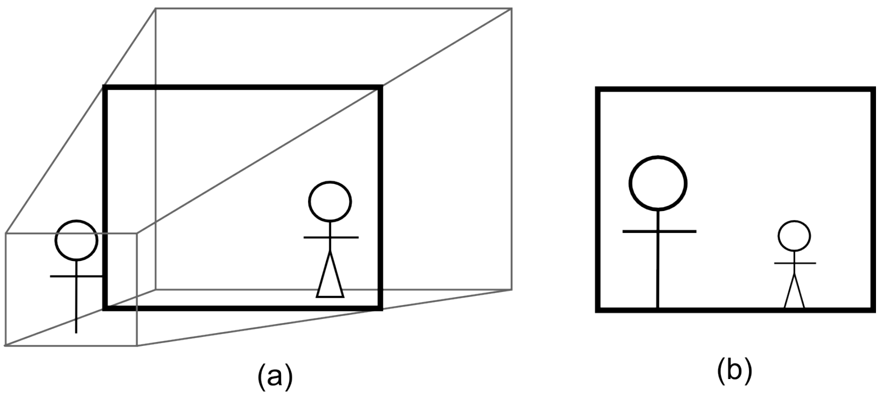


Fig. 2.13 (a) Viewing frustum to implement perspective projection of objects at different distances onto a screen plane. (b) The image projected on the screen plane. Objects in the distance appear smaller than objects close by.

2.2.3 Stereo-pairs using asymmetric frusta

Two asymmetric frusta are used to generate a stereo pair — an image for each eye. These are offset to match the inter-pupillary distance (IPD) of the viewer (Figure 2.14) ensuring that the stereo separation is veridical to real-world viewing conditions. It is important to note that incorrectly configured IPD can feel uncomfortable resolve and cause eye-strain or headache.

The correct eye positions are generally calculated from a tracked point between the user's eyes. A vector that intersects both eyes is calculated — i.e. crossproduct of the 'up' and 'forward' head orientation vectors to obtain a perpendicular vector. The eye positions are translated along this vector as determined by the IPD value.

Using the left-eye as an example Figures 2.15 illustrates where the parameters come from to calculate the extents of the asymmetric frustum (Figure 2.16).

By tracking the head position of the viewer the frustums (for each eye) can be re-calculated in real-time to ensure the accuracy of the image projected onto the screen plane. This provides the motion-parallax cues.

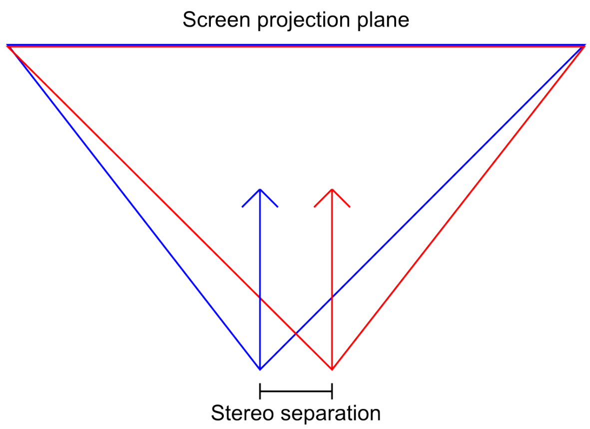


Fig. 2.14 Top-down illustration of asymmetric frusta, one for each eye of the viewer.

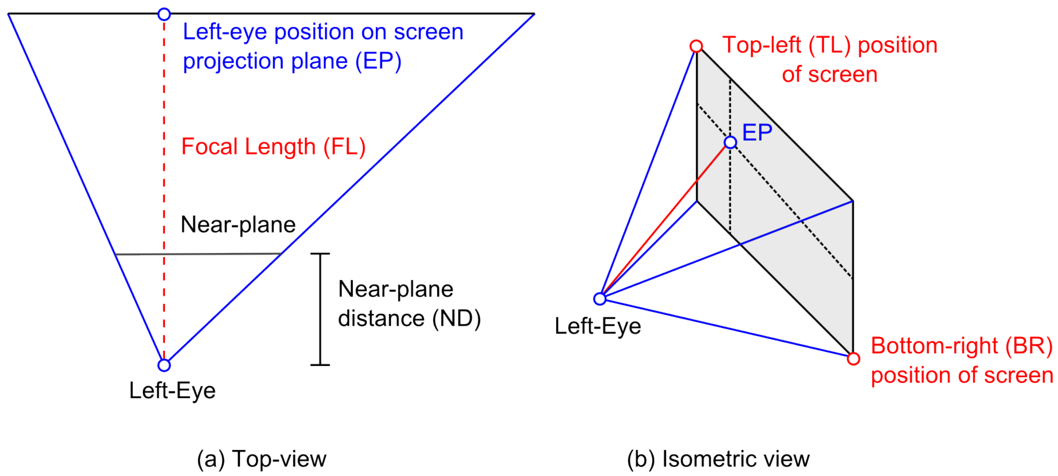


Fig. 2.15 Where the parameters come from to calculate the left-eye asymmetric frustum. (a) The left-eye position at the screen projection plane (EP), the focal length (FL) and the near-plane distance (ND). (b) The top-left (TL) and bottom-right (BR) positions of the screen.

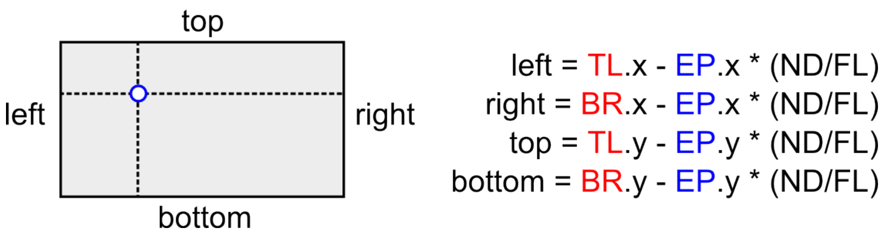


Fig. 2.16 Calculating the extents of the left-eye asymmetric frustum — *left*, *right*, *top* and *bottom*. See Figure 2.15 for the origin of the parameters.

2.2.4 Systems for viewing stereo-pairs

To view a stereo image the left and right images of the stereo-pair must be viewed by the correct eye. HMDs commonly have a separate display for each eye and, therefore, simply display the appropriate image to each eye (Figure 2.17 (a)).

Large screen immersive displays (LSIDs) have 2 broad categories of systems to view stereo, these are 'active' and 'passive'. Passive systems superimpose both the left-eye and right-eye image upon the display surface through polarised filters. The user wears a pair of polarised filter lenses (glasses) so each eye only views the correct image (Figure 2.17 (b)).

Active systems sequentially render the left and right images onto the display surface. The viewer wears a pair of shutter glasses. The shuttering of the glasses is synced to the display output so that when the left image is displayed the right eye is shuttered and when the right image is displayed the left eye is shuttered (Figure 2.17 (c)). Shutter glasses are usually implemented using LCD lenses that have the property of becoming opaque or transparent — i.e shuttered or open.

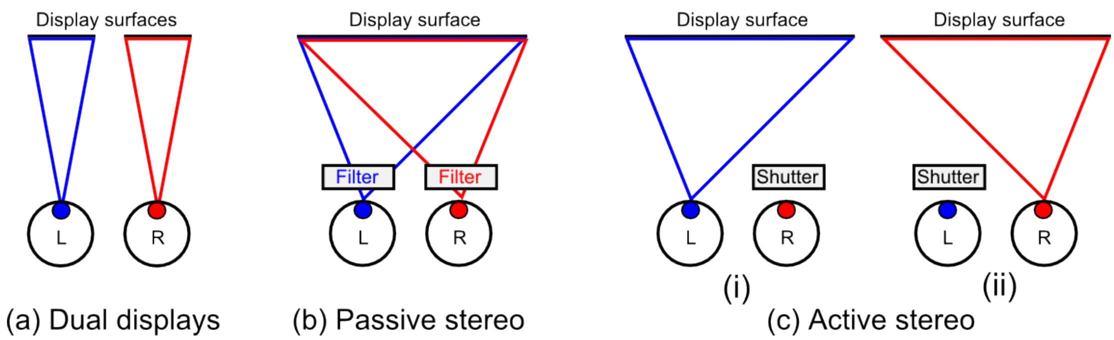


Fig. 2.17 Plan view illustration of stereo-pair viewing systems. The blue triangle represents the left-eye (L) frustum and the red triangle represents the right-eye (R) frustum (a) Duel screens. Each eye has its own display (b) Passive stereo system. Both the left and right images are superimposed on the screen. The glasses filter the images (c) Active stereo system. (c-i) When the left image is displayed the right-eye is shuttered. (c-ii) When the right image is displayed the left-eye is shuttered.

2.3 Identified deficits of depth perception in VEs

The most elementary function or goal of a VE is to generate images to portray a perceivable virtual space that is indistinguishable from the real. Today's VEs achieve this with varying degrees of success. There are numerous known perceptual deficits in the VE that are as a result of limitations in projection and tracking technologies. The causes of some perceptual deficits are more difficult to identify due to limits in the current understanding of human perception in areas such as judging distance or size. The subsections that follow discuss some identified deficits of visual perception in a VE.

2.3.1 Fixed focal distance — accommodation-vergence conflict

Image generation hardware, both head-mounted and wall-projected, only provide a fixed focal distance (Figure 2.18). This creates a conflict between the oculomotor accommodation and convergence. When viewing objects at different depths there are physical movements by the eye; accommodation is the tension exerted on ciliary muscles to focus the lens and convergence is the rotation of the eyes to a focal point. These oculomotor cues are tightly coupled and the discrepancy between them in VE affects depth perception. This is an issue for both a projected VE and a head mounted display (HMD), though there have been some recent breakthrough research with HMD technologies — Liu (2010) developed a prototype HMD with addressable focus cues utilising a liquid lens; it was found that the accommodation cues approximated a real-world viewing condition. Generally, most modern HMDs use lenses forcing the eyes focus at infinity, where the eye muscles are relaxed, again eliminating the ocularmotor cues.

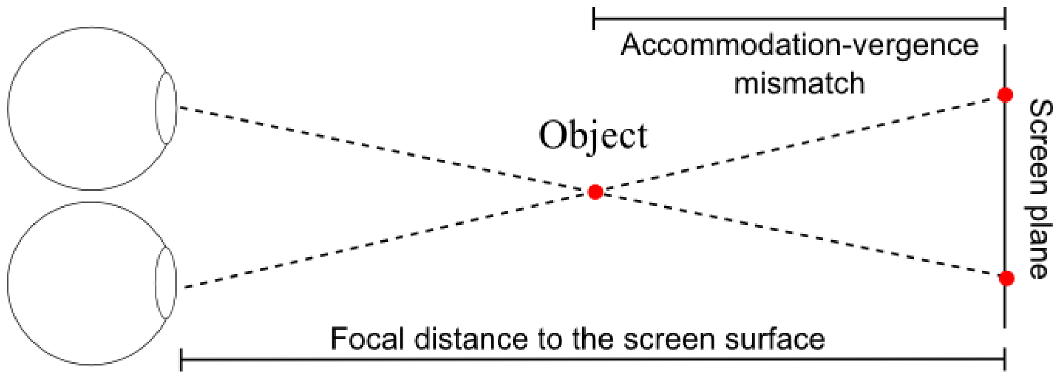


Fig. 2.18 The viewers eyes converge at the object, however, the focal distance is at the screen plane. This creates a conflict between the oculomotor accommodation and convergence cues, known as accommodation-vergence mismatch. This conflict is only mitigated when an object is rendered at the screen.

2.3.2 Impact of visual latency

Head tracking latency causes a display lag that is perceived as the environment swimming about (oscillating) in space, thus degrading the stability of spatial and temporal perceptions. This lag is a result of the end to end latency in the hardware and is refered to as a visual anomaly called ‘oscillopsia’ (Allison et al., 2001). This issue not only affects perception but can also cause what has become known as cybersickness. Cybersickness is defined by Stanney (1995) as a form of motion sickness that occurs as a result of exposure to VEs. The accommodation and convergence discrepancies are uncomfortable for some participants to resolve leading to eyestrain, headache and nausea.

Oscillopsia can be disorientating leading to dizziness and nausea. Meehan et al. (2003) found that visual latency in VE increases heart rate which suggests it causes stress. Any kind of sickness or mental stress within a VE scenario is likely to disrupt performance. Image latency also introduces additional issues for MR applications as the virtuality will not remain correctly overlaid upon the physical world. Papadakis et al. (2011) reported that visible time delay hinders spatial awareness and memory — participant performance of an object recognition task was best in a low latency VE condition.

Jerald et al. (2008) looked at sensitivity to latency at different positions of head yaw and found that users are less likely to notice it at the beginning of a yaw movement than when slowing down or changing direction. The effect of latency is maximised here as the scene is moving against the head movement, thus the user is more sensitive to latency at the edge of head yaws. There are individual differences in thresholds of yaw movement and noticing latency (Jerald and Whitton, 2009).

The lag can be compensated by employing sophisticated algorithms that predict a user head movement. However, Jung et al. (2000) found that the use of these techniques do not offer a a dramatic improvement as they can add noise and overshoot with the image slip in place. Garrett et al. (2002) used a neural network to predict the angular velocity of head movement and found it be more accurate than the extrapolation methods built into hardware at that time. Adelstein et al. (2003) noted that discrimination of image slip is degraded when the direction of head movements are random.

2.3.3 Effect of crosstalk

Visual crosstalk can occur when viewing stereo-pairs. Woods (2011) describes it as the "*the incomplete isolation of the left and right channels so that one image leaks into the other*". To clarify, crosstalk is when the left-image eye leaks into the right-eye, or vice-versa. This creates what is known as the 'ghosting' effect.

Ghosting is the (subjective) perception of crosstalk, it is the appearance of an offset double image or exposure as illustrated by Figure 2.19. This ghosting effect is more noticeable with high contrast scenes. As well as reducing the image quality of a virtual scene, ghosting makes fusing of stereo-pairs difficult, uncomfortable and can potentially cause cyber-sickness.

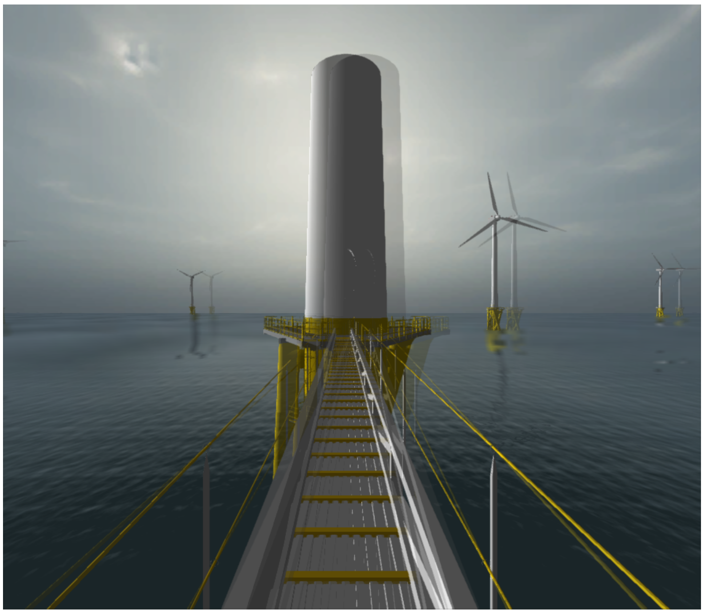


Fig. 2.19 Example of the ghosting effect caused by crosstalk. The scene shows a gangway leading to a partly constructed offshore wind turbine. The image has been manipulated to illustrate the effect of the right eye image of the stereo-pair leaking into the left eye image. The ghosting effect is most evident on the right side of the turbine piece and the rails of the gangway. The ghosting of the distant wind turbines creates a double image effect.

In active stereo systems crosstalk is caused by temporal issue such as poor sync between the display output and the LCD shutter glasses. The LCD lenses can suffer from shuttering persistence i.e. the speed at which LCD lens changes from opaque to clear is too slow. Leakage can occur if the LCD lens is not effective at turning completely opaque. Passive systems can suffer from image leakage if the polarised filters are not accurately aligned. There is also a finite amount image extinction achievable by polarised eye wear. HMDs that use a separate display for each eye will not suffer from crosstalk.

Tsirlin and Allison (2011) conducted an experiment to evaluate the effect of crosstalk on perceived depth from binocular disparity. Using a depth estimation task participants judged the depth between two vertical lines by adjusting a sliding scale. The stimuli were rendered on monitors and displayed using a mirror stereoscopy. The total focal length (viewing distance) was 0.6m away from the participant and the maximum depth between the stimuli was $\sim$7cm. The results showed that crosstalk beyond a 1% level causes a significant decrease in depth perception. Tsirlin and Allison (2012), using a similar apparatus and experiment procedure, investigated perceived depth between objects in 3D photos of natural scenes e.g. such as a kitchen. The results found that perceived depth decreased significantly with only 2–3% crosstalk.

Recently Sohn et al. (2014) investigated crosstalk reduction in 3D stereoscopy displays. They developed a method to displace crosstalk using disparity adjustment to reduce crosstalk in important regions of a scene. However, the ambition of the study is to improve image quality, the impact of the disparity adjustment upon depth perception was not considered.

2.3.4 Empirical studies of distance estimation in VEs

In the real world people are very accurate at judging distances within 20m. The review by Loomis and Knapp (2003) shows that studies in early VEs observed that perception of distance is distorted and that user judgement of egocentric distance in a VE is significantly shorter than the real world. This became known as spatial compression and the body of work that follows attempted to unravel it. Much of the work has focused on medium-field distances ($\sim$2m to $\sim$30m) — what Cutting and Vishton (1995) classified as 'action' space.

The majority of studies utilise HMDs to measured participant perception of distance via direct or blind walking tasks. In a blind walking task experiment a participant is exposed to a real or virtual space, indicating a target on the floor. They were then asked to walk to a target while blindfolded. Participants could do this with great accuracy in the real world, but would always walk less distance in the virtual condition. Although most of the work in distance perception relates to HMD walking based tasks, Sahm et al. (2005) investigated the additional action based task of throwing, comparing it to walking. The results showed that spatial compression occurred similarly in both tasks indicating that the cause of the spatial compression is task independent.

Research was also carried out to investigate if spatial compression is caused by the rendering of the graphics, HMD hardware, large screen immersive display (LSID) hardware and presence in a VE.

Willemsen and Gooch (2002) thought likely factors causing spatial compression were missing information in the rendering of the virtual scenes and artefacts of the display technology. They compared and found that the difference between two types of VE rendering, panoramic photo and traditional polygon rendering, were small. For the distances of 2–5m that were investigated, spatial compression was present with both rendering techniques suggesting that the HMD itself could be the source of issue. Thompson et al. (2004) investigated the quality of rendering as a probable cause of spatial compression using three rendering qualities but found it made no difference. They concluded that the photo-realistic improvements of texturing and illumination did not aid judgements in a VE. However, contrary to the conclusions made by Thompson et al. (2004), Kunz et al. (2009) found that the quality of graphics made a difference for verbal reporting tasks, suggesting that the cues required for depth perception depend on the context of the task.

Willemsen et al. (2008) manipulated the stereo viewing conditions — non-calibrated and calibrated inter-pupillary distance (IPD), binocular and monocular viewing. They eliminated the accommodation and vergence cue conflict using monocular viewing and found that none of the factors improved overall performance the problem, concluding that inaccuracies in stereo-viewing conditions were not likely the source. In Willemsen et al. (2009) the mechanical aspects of the HMD, the FOV, mass and moments of inertia, were investigated and it was concluded that the spatial compression effect is not a result of an interaction between HMD mechanics and the requirement of walking. Kuhl et al. (2009) stressed the importance of HMD calibration in studies and investigated the pitching of the virtual world due to mismatch of sensor and optical display, and pin-cushion distortions. However, correcting these made no significant effect on spatial compression. A real world study by Knapp and Loomis (2004) simulated the limited FOV of an HMD and concluded that it could not be the cause of spatial compression in a VE. They speculated that may be accommodation-vergence mismatch, the limited dynamic range and spatial resolution of display and rendering fidelity that causes spatial compression. Another real world study by Creem-Regehr et al. (2005) examined horizontal FOV and binocular viewing restrictions and concluded that it could not largely contribute to spatial compression. They also found that a restricted head movement increased spatial compression leading to speculation of a need to scan the nearby environment before judging distances.

Using a low latency and high fidelity HMD, Interrante et al. (2006), found that within an exact virtual replica of the space currently occupied by a user, spatial compression is not significant, concluding that perhaps participants are more immersed in virtual space that has already been experienced and accepted in the real world. As a result the group formed a ‘visual calibration hypothesis’ in Interrante et al. (2008), stating that participants rely on information gained from exposure to the real environment to calibrate judgements. However, the experiment conducted proved this to be false when they found that spatial compression was not mediated by a user's initial exposure to the exact real environment. However, they made the interesting observation that over successive virtual trials the participant walked for longer, suggesting a general effect caused by time spent in the VE.

Mohler et al. (2008) based their study on the questions of, if awareness of one's body could provide a metric for scaling of absolute dimensions of space and whether the presence of the body in the VE could serve to ground or situate the user in the VE, acting as a frame of reference. They conducted an experiment with a full body avatar, whose position and orientation were tracked and rendered, and found that spatial compression was reduced. A similar study by Ries et al. (2008) had similar results concluding that an avatar facilitates a stronger sense of presence in the environment and that the spatial compression phenomenon may have been caused by cognitive dissonance caused by VEs. A further study by Mohler et al. (2010) found that participants that experienced a fully articulated and tracked co-located avatar, that stood 3m ahead of them, made more accurate judgements for distances between 4–6m than those that did not. They further conclude that their results introduce new questions of the nature of self representation in VE and its effect on spatial perception. However, it is interesting to note that Creem-Regehr et al. (2005) found that being able to see feet or body did not matter for distance perception in the real world.

Studies of spatial compression in LSIDs cannot facilitate visually guided action based distance measures of blind walking or throwing — due to the physical space constraint. In response Klein et al. (2009) proposed and investigated distance perception using three measurement protocols suitable for LSIDs. The study were conducted between 2–15m, with imagined walking, verbal estimation and triangulated blind walking, and found that the results for the imagined walking and verbal estimation were equivalent to real world estimations. However, triangulated blind walking, which involved turning 90 degrees, walking 2.5m metres then pointing at the distance. However, the results of the triangular walking task was only accurate in the real condition suggesting that knowledge of the room geometry interfered with the result.

Murgia and Sharkey (2009) devised a judgement task to see if size constancy could alleviate the spatial compression. They judged the distances to virtual replicas of familiar physical objects, placed at distances up to 3m by placing a virtual marker where they perceived the object to have been. It was found that familiar objects did not improve spatial compression. More recently, Piryankova et al. (2013) studied motion-parallax and stereo cues in LSIDs using a verbal reporting measure and found an underestimation of depth in comparison to nearly accurate estimates in the real world.

Near-field distance perception (within $\sim$2m) — what Cutting and Vishton (1995) classified as ‘personal’ space — in VEs has received less research attention. Napieralski et al. (2011) reported significant underestimation when comparing near-field distance perception between the real world and a HMD VE. Investigating depth compression in LSID, Piryankova et al. (2013) found it was present in the near-field, but stereoscopic depth cues mitigated it somewhat.

Ellis and Menges (1998) investigated near-field distance perception to AR objects using a blind-reaching task. Overestimation was observed when an AR object is at the same distance of a background surface i.e. superimposed upon it. When a surface was nearer to the participant than the AR object — occluding — it resulted in an underestimation of distance. It was noted that older participants struggled to localise AR objects at close distances, it was concluded that this was due to an inability to accommodate at such short focal lengths and instead relying on disparity cues.

Singh et al. (2010) replicated the apparatus of Ellis and Menges (1998) and observed distance underestimation to near-field AR objects. It was found that participant’s would perceive the AR objects to be at the same distance of a salient occluding surface if there was a feasible association between them.

2.3.5 Empirical studies of size estimation in VEs

The inability to judge distance accurately within a VE is a severe restriction and it may also have perceptual consequences on other abilities such as judging the correct size of an object or more specifically size constancy scaling. Size constancy scaling was first described by Descartes in 1637, (modern translation by Cottingham et al. (1988)), as the tendency to perceive an object as having the same size despite changes in visual angle. Visual angle is the angle subtended by an object at the observer’s eye. It varies with object distance, and is used as a measure of the projected image size on the retina. The size constancy scaling mechanism compensates for the changes in retinal size so objects do not appear to grow or shrink at different distances.

Traditionally, size constancy scaling was thought to depend on depth information (depth cues). Studies by Holway and Boring (1941) demonstrated that removing depth cues made the participant rely purely on the optics of the situation i.e. visual angle. Kenyon et al. (2007) recently performed an experiment demonstrating to what extent size constancy scaling could be experienced in the CAVE. Elner (2007) conducted a study in the Hull immersive visualization environment (HIVE) that yielded a similar result. The VE studies by both Kenyon et al. (2007) and Elner (2007) used elements from the Thouless and Holway experiments confirming that size constancy mechanisms do occur in the VE and affirmed the previous belief that the depth cues are paramount for it. The apparent importance of depth cues for distance perception can support a claim that incorrect distance judgement may affect size judgement.

Size and distance perception, though parallel processes, may be less coupled than previously thought. Haber and Levin (1989) argues that not only do distance and size perceptions meet separate perceptual demands, they also depend on largely separate non overlapping sources of information. Haber and Levin (2001) make a strong argument that size perception is based upon memory and familiarity and any concurrent real time processes such as distance perception are irrelevant. Their hypothesis proved true for objects that had a standard size e.g. a football or cup. However they could not explain how participants could correctly judge the size of familiar objects at distance that could vary in size.

Elner (2007) made an anecdotal observation that when stereoscopy was turned off within the HIVE, uniformly scaling an object from its centre would be incorrectly perceived as movement towards or away from a participant, in other words size constancy scaling occurred inappropriately. The object in question was not casting a shadow on the environment — a dominant cue for judging distance. When a shadow cue was introduced a participant would correctly perceive the object scaling as a change of size and not distance. This is interesting because the removal of key depth cues actually promoted size constancy scaling in spite of the previously discussed studies indicating that removal of depth cues tends to result in participants depending on the optics of the situation i.e. the subtended visual angle. Elner (2007) speculated that this could be a cognitive reaction as familiar real-world objects do not uniformly grow or shrink unless they are moving towards or away from us. The object in question did have a familiar real world counterpart so this would seem to support Haber and Levin (2001) theory.

According to Eysenck and Keane (2005) the current general consensus is that size constancy scaling depends both on perceived distance and familiarity of the object in question, though there is no coherent account of how. Elner (2007) observation could suggest that knowledge was secondary to the concurrent processing of depth information. Both distance and size perceptions are essential components of spatial awareness and ability. The identified deficit in spatial distance judgement is an established one, but whether this does in fact impact upon perception of size needs to be established.

2.4 Defining the scope of the research

The term VE is generally used to broadly label all configurations of VE hardware with no regard to the classification of the virtual spaces they create or the enabling technologies. For convenience the term VE is used in this thesis in the broadest sense. Although non-visual elements — haptic and auditory — can be important for conveying a sense of presence and are increasingly found in VEs this research is focused on the visual modalities of the synthetic space.

This section defines these different virtual spaces and the visual display modalities that implement them to clearly establish the scope of the experimental work.

2.4.1 Classifications of VE spaces

Ellis (1994) provides a useful definition of a VE as "*[an] interactive, virtual image display enhanced by special processing and by non-visual display modalities, such as auditory and haptic [touch], to convince users they are immersed in a* ***synthetic space***". However, the synthetic spaces are not necessarily fully virtual as they can merge objects from both the physical and virtual realities within a single shared space. Milgram and Kishino (1994) refer to these shared spaces as mixed reality (MR) and describe them as a subset of virtual reality that sits along a virtuality continuum (Figure 2.20). At the two extremities of the continuum are the real or physical environment and the VE, and between these is MR. Moving along the continuum from the physical world towards the virtual world, the physical world is augmented with virtual objects, which is known as augmented reality. Thomas (2012) defines AR as "*the registration of computer-generated graphical information over a user's view of the physical world*". Moving further along the continuum from AR towards the virtual world sees the amount of virtual objects increase until the majority of the environment becomes virtual, but retains some real objects. This is the opposite of AR and is termed an augmented virtuality (AV). At the opposite extremity to the real environment is the VE, where the space is entirely composed of virtual objects.

Although Milgram and Kishino (1994) provide a useful classification system, the virtuality continuum does not address the localisation of the VE space. This is addressed by Benford et al. (1996), a broad classification of shared spaces (mixed reality) according to two dimensions: transportation and artificiality. As seen in Figure 2.21, the artificiality dimension is akin to the virtuality continuum, whilst the transportation dimension concerns "*the extent to which a shared space introduces new information into its users' local spaces versus the extent to which it allows them to enter remote spaces*" (Benford et al., 1996).

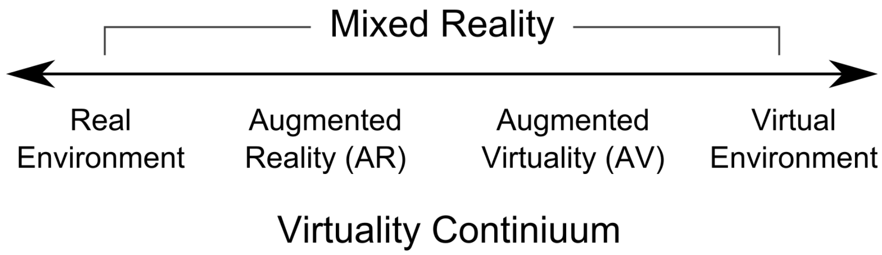


Fig. 2.20 Virtuality continuum recreated from Milgram and Kishino (1994).

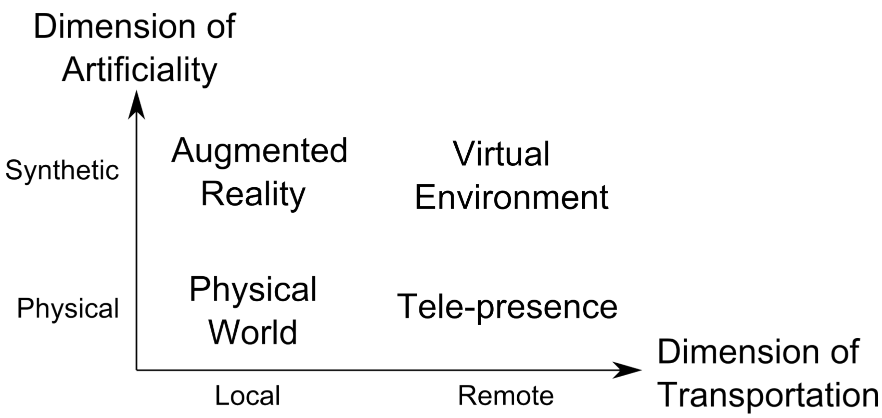


Fig. 2.21 Broad classification of shared spaces according to transportation and artificiality adapted from Benford et al. (1998).

Using these two systems of classification the terms VE, AR and AV are considered within the context of this research as follows. VE is a completely remote synthetic space where the user has no perceptual frame of reference within the 'local' space they physically occupy; meaning they are occupying the virtual space. AR is a MR environment that augments the user's local space with virtual objects while the user is still perceptually anchored to the the physical space they occupy. AV is MR environment where the user is within a 'remote' synthetic space, with aspects of physical world present where they exist within the virtual space.

2.4.2 Visual display modalities

Virtual image display technologies generally fall into one of three categories: head mounted display (HMD), hand-held device and display surfaces.

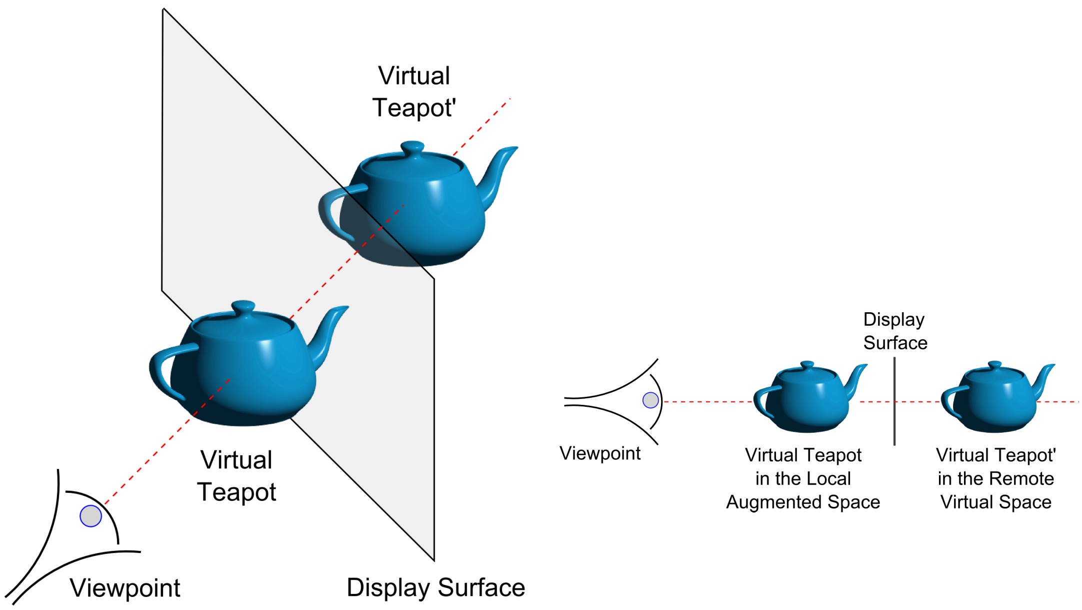
(a) (b) (c)

Fig. 2.22 Three examples of VE display technologies (a) Display surface — four sided active stereo projection based VE at the University of Hull. The user wears LCD shutter glasses and is head tracked (b) HMD — shows the OculusVR (2014) Occulus Rift. Note the screen behind the subject displays the images that each eye views (c) Handheld devices — a tablet using IKEA (2013, 2014) augmented reality furniture application. The devices use a camera to grab an image of the real room and overlay a piece of virtual furniture.

A HMD simulates depth by displaying a stereo image to each eye — binocular stereo. This visually blocks the user's local physical environment. The user's physical head orientation and position can be motion tracked allowing them to visually explore the VE. There are also HMDs that allows the view of both the real and virtual environment at once — AR. The first variant is 'optical see-through' which displays images by reflecting them onto a transparent medium, allowing for an unhindered view of the physical environment. The other type is known as 'video pass-through' uses a video camera to capture an image of the physical environment an displays it on a opaque medium. Such AR devices have been applied to medical surgery (Rolland and Fuchs, 2000).

Hand-held devices (Figure 2.22 (c)) such as tablets and smart phones are often used for AR using the in-built camera for video pass-through. The software that drives these applications uses image processing techniques to calculate the device camera position for tracking and devices with auto-stereoscopy screens can project 3D imagery. However, given the limited display size and the inability to project in negative stereo (screen occlusion) means that these devices are often limited to be used as a portal or window into a VE dimension.

Display surfaces (Figure 2.22 (a)) normally consist of tiled displays or are projection-based such as the CAVE (Cruz-Neira et al., 1992). Such set-ups can provide a combination of both AR and VE at once. The space generated beyond the screen is virtual, however, the physical space in front is augmented (Figure 2.23).



(a) Isometric view (b) Orthographic side-view

Fig. 2.23 Local and remote virtual spaces of display surface VEs. The *virtual teapot* that appears to be in front of the display surface is within the viewers local physical space. It is augmenting the physical space with a virtual object so it can be considered as AR. The *virtual teapot'* that appears beyond the display surface occupies a remote synthetic space, that is impossible for the user to occupy.

2.4.3 Scope of experimental work

The research scope of the experimental work is large surface display VE. However, the term mixed reality is a more appropriate description, as the experimental work will explore the AR local space in front of the screen.

2.5 Summary

As required by the first objective, the research background in this chapter has identified the deficit of depth perception in a VE and discussed the body of work pertaining to it. Furthermore, the type of VE to be studied has been identified as to be a large surface display VE.

Chapter 3

Study of Phenomenal Regression

"*OK, one last time. These are small, but the ones out there are far away. Small... Far away... Ah forget it.*"
— Father Ted Crilly

This chapter investigates the psychological trait phenomenal regression (PR) to fulfil objective [ii] of the research aim. PR is potentially a useful metric that has yet to be applied to the study of visual perception in a VE. PR has properties that make it an immediately attractive candidate for this research. Whereas other groups have used skill-based measures PR instead offers a 'direct measure' of subjective experience in a VE. PR is not mediated by experience, skill or previous training meaning that subjects cannot get better at it with practice. PR measures an innate, involuntary and unmodifiable response to visual stimuli, it cannot be manipulated. Thouless (1932) demonstrated that only artificial reduction of depth cues changes PR. We propose that it can be used to assess veridical perception in a VE by measuring and observing the difference in users' PR between the real world and VE.

This chapter provides an informed restating of PR, by revisiting the original research by Thouless and conducting exploratory work. The original papers by Thouless (1931a,b, 1932, 1933) are challenging to read and difficult to internalise and understand. Written in the 1930s, the research is reported in a language and style that today seems quaint. It is lacking in detailed apparatus set-up and the chronology is jumbled with clues and useful details omitted from earlier papers but mentioned in passing in later. As such there are aspects of PR that are not immediately understood from reading the original papers. From examining the original research and conducting exploratory work, we re-describe PR with an informed understanding.

This chapter is structured as follows: section 3.1 describes PR, its numerical value TR and the experimental methodology to measure it. It also highlights the internal and external psychophysical factors that influence TR. Section 3.2 retraces the steps that led to the current level of understanding of PR by highlighting the difficulties and then discussing the exploratory experiments that helped clarify it. Based upon the original Thouless papers and the exploratory experiments section 3.3 lists the presumptive properties of PR.

3.1 Explanation of PR

Thouless (1931a) found that participants make a systematic error when asked to match the perceived characteristics of two stimuli. It was found that the perceived characters of objects in the world do not match those of retinal stimulation, but are influenced by the character of the real physical attributes of the object.

This section describes PR and the methodology to measure it, by using the character of shape as an initial example. This is followed by an example of PR using the character of size, which is a slightly more complex calculation as it uses the two linked quantities of size and distance. Thouless (1932) also investigated PR of brightness and hue (colour); however, these are not discussed as they are not attributes that are of interest to this research.

Finally the internal and external psychological factors that influence PR are discussed.

3.1.1 PR of shape

Thouless (1931a) conducted an experiment investigating perspective shape matching. The participant was presented with an inclined disc and asked to pick, from a series of ellipses, one that matches the shape of the disc's perspective projection, which is the visual angle of the shape subtended at the eye. The participant was not asked to judge the real shape of the inclined disc, i.e a circle, but instead what they saw i.e. an ellipse. It was found that participants did not choose an ellipse that matches the perspective shape. They chose an ellipse that was somewhere between the perspective shape and the actual, circular, shape (Figure 3.1). The shape of perception is the phenomenal (P) shape, the actual shape of the object is the real (R) and the shape of R projected onto the retina of the eye is the stimulus (S). Thouless' results found that P shows tendency from S towards R and referred to this as phenomenal regression to the real object.

PR is measured using a normalised measure. The general formula to normalise a value

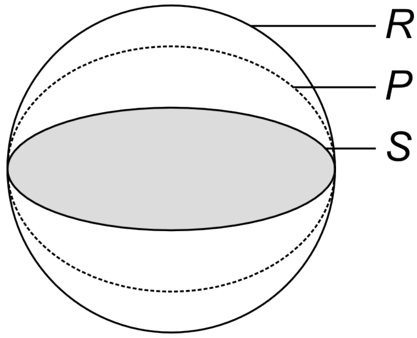


Fig. 3.1 The real (*R*) shape of the disc, as indicated by the solid line, is a circle. Inclined to the line-of-sight of the viewer, the disc casts the stimulus (*S*) shape (shaded ellipse) onto their retina. It is found that the perceived shape or phenomenal (*P*) shape (dotted line) actually lies somewhere between *R* and *S*. [Redrawn from Thouless (1931a) Fig. 2, p.342].

within a range is:

$$x' = \frac{x - min(x)}{max(x) - min(x)} \tag{3.1}$$

where x is the original value and x' is the normalised value. PR is thus calculated as:

$$\frac{P - S}{R - S} \tag{3.2}$$

where S, P and R are the stimulus, phenomenal and real shapes of the disc. The difference between P and S, the amount of regression, is normalised over a baseline value that measures the total difference between R and S. Therefore PR is P normalised to have values between zero and unity. The equation in this form is commonly known as the Brunswik ratio (Sedgwick, 1986) — named after the psychologist that devised it. Thouless (1931a) expressed PR using a variation of 3.2:

$$\frac{\log(P) - \log(S)}{\log(R) - \log(S)} \tag{3.3}$$

Thouless (1931a) called this measure the index of phenomenal regression (IPR). However, today it is generally known as the Thouless ratio (TR). Taking logarithms of the various terms in the Brunswik ratio confers an advantage that will be described later. If a subject performs a perfect perspective match, P is reported as S and TR is therefore zero:

$$\frac{\log(S) - \log(S)}{\log(R) - \log(S)} = 0 \tag{3.4}$$

If a subject performs a perfect objective match, P is reported as R and TR is unity:

$$\frac{\log(R)-\log(S)}{\log(R)-\log(S)}=1 \tag{3.5}$$

The values for S, P and R for the shape matching experiment are calculated as the ratios of the major and minor axes of their elliptical shapes (Figure 3.2). S is the value of the axis

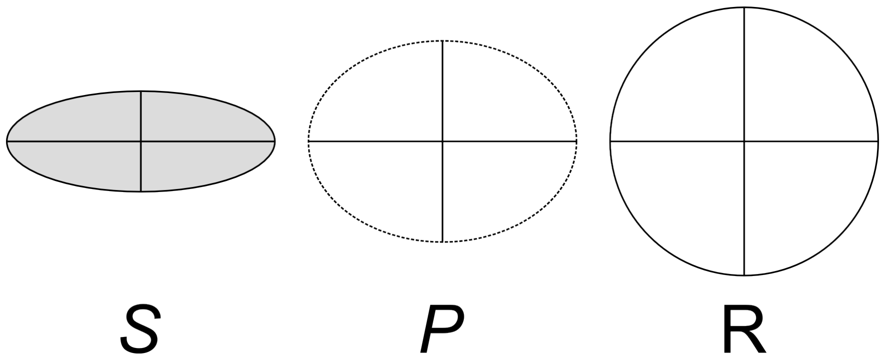


Fig. 3.2 The ratio of major and minor axes of the elliptical shapes of R, S and P. Since the disc in the shape experiment is circular, R is always unity.

ratio of the perspective ellipse projected onto the observer's eye. Although Thouless (1931a) took a photograph at the observer's eye position to obtain a physical representation of the ellipse from which the ratios could be measured this can be calculated using trigonometry. P is the axis ratio of the disc chosen by the participant. R is the real shape of inclined disc, i.e. a circle, and therefore has axis ratio equal to 1. TR of the character of shape is therefore:

$$\frac{\log P-\log S}{-\log S} \tag{3.6}$$

3.1.2 PR of size-distance

Thouless (1931a, 1932) also investigated PR of the size character of an object. Participants were presented with 2 circular discs of different sizes d and D at respective distances l and L (Figure 3.3). The distance, l, of d is adjusted until it is reported that its perspective size matches that of D. The adjustable disc is referred to as the 'response' and the comparison disc as the 'standard'. As with the character of shape, PR of size is found to be a compromise between the stimulus and real characters (Figure 3.4).

In the size-distance experiment the stimulus is the size that the standard D would have to assume at the distance, l, of the response d, in order to appear as it does at L i.e. Dl/L

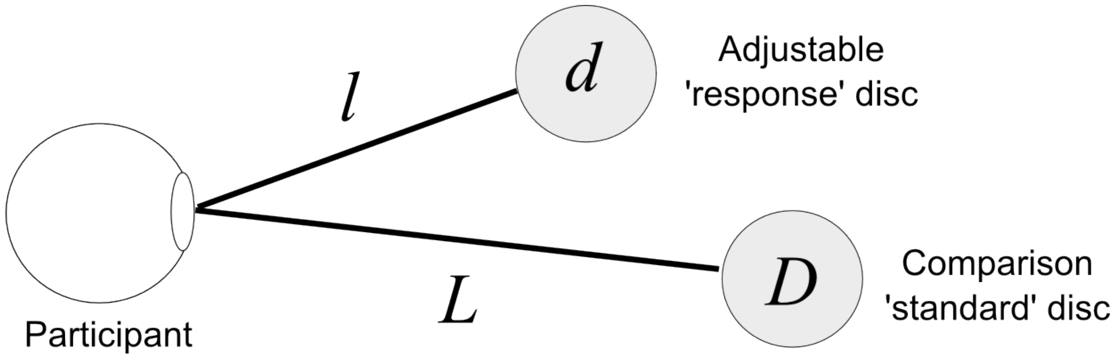


Fig. 3.3 The arrangement of discs in Thouless' size-distance experiment. A disc with diameter D is placed at distance L and a disc with diameter d is placed at distance l. The discs are arranged such that they do not overlap in the viewer's field of view.

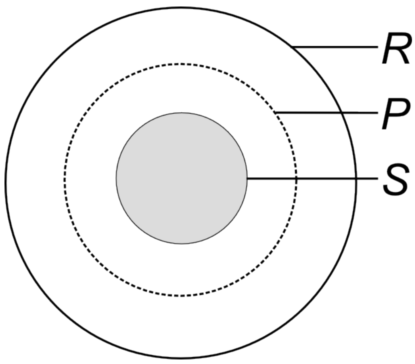


Fig. 3.4 The real(R) size of the disc (solid line), the perspective size of the stimulus (S) cast onto the retina (shaded). It is found that the perceived or phenomenal (P) size actually lies somewhere between R and S (dotted line). [Redrawn from Thouless (1931a) Fig. 5, p.352].

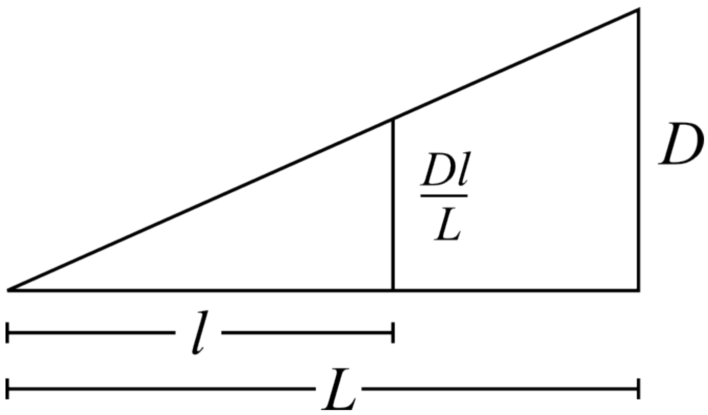


Fig. 3.5 The size at l of the standard disc D at L is calculated using similar triangles.

(Figure 3.5). Adjusting the response disc d to match the standard disc D leads to

$$TR = \frac{log(d) - \log(\frac{Dl}{L})}{\log(D) - \log(\frac{Dl}{L})} \tag{3.7}$$

This equation can be rearranged by adding $log(d) - log(d)$ to both the numerator and the denominator, to result in:

$$TR = \frac{-\log(\frac{Dl}{dL})}{\log(\frac{D}{d}) - \log(\frac{Dl}{dL})} \tag{3.8}$$

In contrast to PR of shape, it can be seen that we now handle two linked quantities, size and distance, to form the ratios. The ratios used are the ratio of the two disc sizes, rather than the axes of one disc, and the ratio of their perspective sizes. The baseline of the equation is, therefore, now the difference between the real size ratios and the perspective size ratios.

The properties of logarithms mean that it does not matter which disc in the experiment is adjusted (the 'response' disc) and which is the used for comparison (the 'standard' disc). In other words, the response and standard are interchangeable, since rearranging 3.8 gives:

$$TR = \frac{-\log(\frac{dL}{Dl})}{\log(\frac{d}{D}) - \log(\frac{dL}{Dl})} \tag{3.9}$$

This is not the case if the Brunswik ratio is used, hence why it is inferior. To summarise, TR of the character of size is:

$$TR = \frac{-\log(S)}{\log(R) - \log(S)} \tag{3.10}$$

where S and R are the stimulus and real characters of size of the discs. The character P is matched to phenomenal equality i.e. $P = 1$ so $log(P) = 0$.

Thouless (1931a) reported problems in administering the movement of the response disc and said that with hindsight it would be better to have adjusted its diameter instead of changing its distance from the participant. This also avoids pitfalls identified by others: Gottheil and Bitterman (1951) criticised Thouless' experiment methodology stating that PR cannot ever be complete ($TR = 1$) when the response and standard disc used are of different physical sizes. McDonald (1962) found an artefact of the Brunswik ratio equation, reporting that variations in distances l and L change its value leading to incorrect interpretations of results. We assume that this artefact is present in the Thouless ratio, which suggests that results between different experiments are only comparable if the same distances of l and L

have been used. The solution is thus to keep distances l and L constant and instead ask the participant to adjust the diameter of the response disc until its perspective size matches the standard.

3.1.3 Factors that influence PR

PR is subjective and shows individual differences ranging across the TR scale. It is also found to be influenced by both internal and external factors, where the internal factors are personal physiological and psychological traits of the observer and the external factors refer to the visual stimulation from the environment.

Thouless (1933) investigated a number of traits of an individual's psychology and how each influences PR. It was found that individuals with higher intelligence showed less tendency to regress towards the real characters of objects. Females showed TR values nearer to unity than males. Although not statistically significant, it was observed that PR increases with age. This is certainly feasible given that the accompanying physiological changes, like the hardening of the lens of the eye with age, reduce the ability to accommodate (focus). Participants with art training had lower TR than those without. Artists are trained to draw in perspective and have, therefore, practised the ability to judge the characters of objects by retinal image. The research also tested nationality and observed that subjects from India showed statistically much higher PR than British.

Thouless (1933) found that an individual that showed large PR of size would also show large PR of shape; however, he found no such no link to the character of brightness.

Regression towards the real character cannot happen unless there is some awareness of the real character. Thouless (1932) concluded that this awareness is not dependent on previous knowledge of the object, but on the presence of perceptual indicators i.e. for the size matching this would be the relative distances between discs. In a monocular viewing condition a participant's PR of shape is reduced, suggesting that PR is sensitive to changes in stereo cues. It was found that despite the reduction of PR, it is never to the extent that the phenomenon matches the stimulus. Not being able to totally eliminate PR is possibly the result of the internal observer traits.

To summarise, PR differs between individuals with identified internal and external factors. As discussed in Chapter 2, there are numerous theoretical models on how the brain forms a perception of the world and the objects within it. Although this in itself is not of interest to this study, we do care about the inputs and the output. By thinking of the brain as a black box (Figure 3.6) we can bypass the question of 'how' and focus on which inputs influence the output — phenomenal perception.

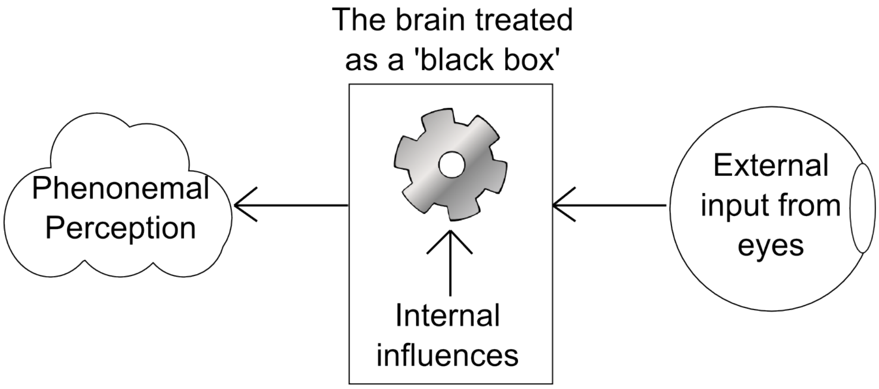


Fig. 3.6 By treating the brain as a 'black box' we side-step the debate of how phenomenal perception is formed. We are only concerned with the internal and external factors that influence phenomenal perception. The internal factors are traits of individual psychology, such as intelligence, age and sex. The external factors are the perceptual cues provided by the real object.

3.2 Retracing steps to understanding PR

Given the apparent simplicity of the effect we set out to measure, it is worthwhile to reflect on the difficulties faced in attaining this level of understanding and to retrace the steps taken to achieve it.

From the outset there were aspects of measuring and calculating TR that were easy to understand. For instance, the experiments of shape measure a single quantity, measured as the ratios of the major and minor axes of the discs which makes it easy to derive the equation for TR as well as to sketch and visualise. Understanding the size-distance phenomenon was less straightforward as it was not immediately evident what constituted the ratios. Initially the equation (1.9) as on p353 of the 1931 paper was used as a recipe while the full understanding of its derivation came about much later.

As mentioned, the original research is difficult to read, leading to difficulties understanding aspects of PR of shape and size. Some of the complexity is due to the writing style and sparsity of diagrams; for example, at inception there was no diagram of apparatus arrangement for the size study. Key misunderstandings arose from the lack of clarity of the phenomenal character P as well as the aim of the matching task given to participants. Thouless' instructions did not shed any light on the intent of the matching tasks. Although we are now certain that perspective matching was the goal of the instruction, it was initially thought that the intention of the task was to match the real physical aspects of shape and size (to perform objective matching).

Section 3.2.1 summarises the exploratory experiments carried out to help clarify the misunderstandings mentioned above. Although it was designed and conducted formally, the manner in which the participants were instructed was found to be flawed as detailed in section 3.2.2. This is followed by a discussion of the difficulty in understanding the terminology used in Thouless' original papers in section 3.2.3. The final section 3.2.4 describes how the understanding of the phenomenal character, the aims of the task and therefore the intentions of Thouless' instructions were achieved. Section 3.2.5 is a summary of the research output of the exploratory work.

3.2.1 Summary of pilot experiments

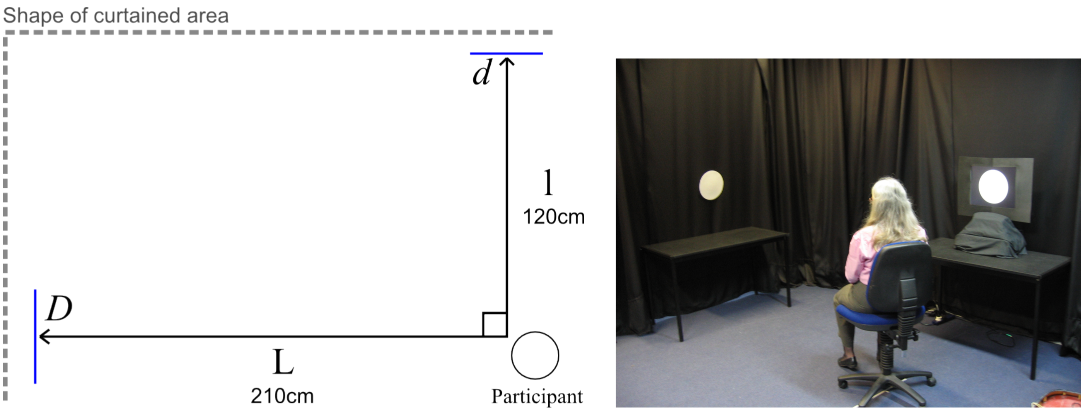
The aim of the experiments was to observe PR and whether it occurs in a VE. Additionally it was felt that reproducing Thouless' experiments would clarify the understanding of the methodology for measuring TR. Thouless' size-matching experiment was chosen because, unlike shape, it has an additional linked quantity of distance, in order to obtain ratios of stimulus sizes. This means there is some synergy between previous VE studies on distance perception.

The experiments were designed to measure TR of participants in two environment conditions, the real and virtual, summarised in Figures 3.7 (a) (b) and 3.8 (a) (b). Although efforts were made in the VE condition to immerse the participant in a space similar to that of the real condition, it is not identical. In the VE the angle between the discs is restricted due to the physical constraints of the display set-up. The experiment procedure was the same in both real and VE conditions, the participant was instructed to adjust the size of response disc d until the size matched the standard disc D.

The results are summarised in Figure 3.9. There is no statistically significant effect or correlation between the real and virtual conditions. The bar chart indicates a range of TR values hinting at the individual (subjective) nature of PR. However, the participant instructions were ambiguous and could have been interpreted to mean either a perspective or objective match and thus the results are not verifiable, as detailed in the section that follows.

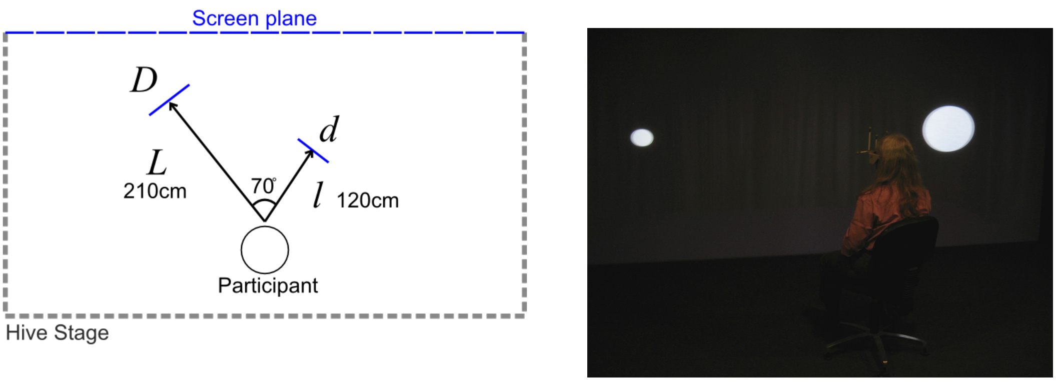
3.2.2 Difficulty using Thouless' instructions

Participants were instructed by following a procedure similar to the one used by Thouless (1932). The paper provides the wording of Thouless' instruction which we used verbatim



(a) (b)

Fig. 3.7 (a) Apparatus arrangement of real world pilot experiment. The response disc *d* and the standard disc *D* are arranged at distances *l* and *L* from the participant. The discs are arranged at a 90 degree angle and the distances were chosen to mimic those of the original Thouless (1931a) experiment. (b) Photo of the real word pilot experiment. The response disc is displayed in computer graphics on an LCD screen and is adjusted using a wheel on a wireless computer mouse. The standard discs are made from felt and are pinned to the curtain. Four sizes of standard disc are tested: 15cm, 19cm, 23cm and 27cm. These are similar sizes those of the original Thouless (1931a) experiment, also 27cm was the limit of what could be displayed on the LCD monitor.



(a) (b)

Fig. 3.8 (a) Apparatus arrangement of the VE pilot experiment. The response disc *d* and the standard disc *D* are arranged at distances *l* and *L* from the participant. (b) Photo of the VE pilot experiment. Both discs are virtual, generated with stereo and motion parallax cues and are perceived to be at *l* and *L*. The angle between the discs is 70°.

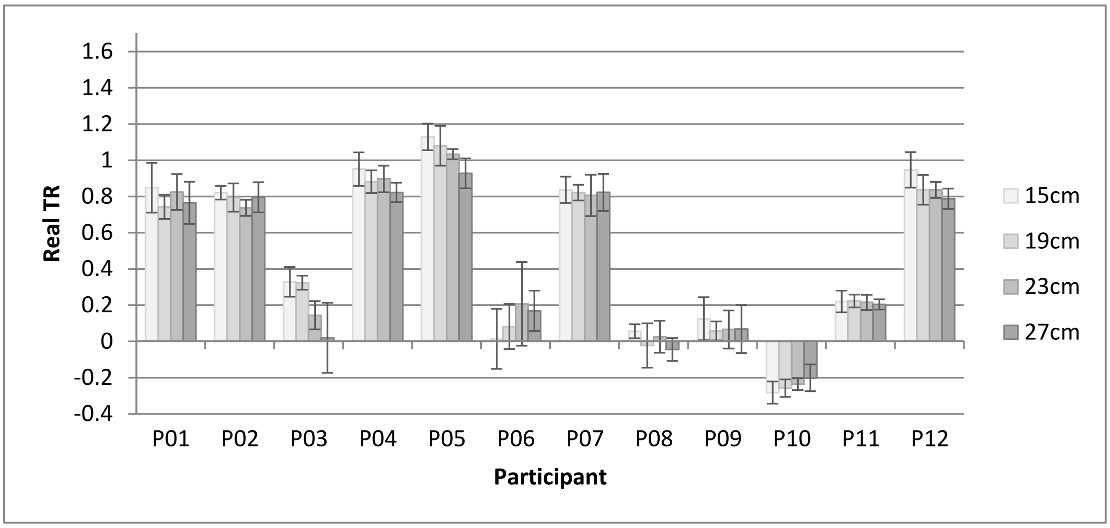
in the experiments hoping to reproduce his findings. This was a mistake, as it left us ill-equipped to deal with that fact that participants struggled with the instructions. There was a sense that the PR of size was about objective matching i.e. size constancy, but we weren't certain.

A 27cm disc made of felt was fixed to be presented to the participant facing them at their eye level at a distance of 1.22m away. A second smaller 15cm disc is held up next to it and the participant is asked the following question *"This is a large disc and this is a small one. This one looks smaller than the other, doesn't it?"*. The question makes the point that when the discs are placed side-by-side at the same distance the smaller disc is in fact physically smaller than the larger disc. In other words it makes clear the real size of the objects. The experimenter then slowly brings the smaller disc towards the participants eyes, being careful not to occlude the larger disc, saying *"As I bring it nearer to your eyes it looks bigger and bigger until it finally looks bigger than the other one"*. We now know that this statement is highlighting the obvious fact that as the smaller disc gets closer to the participant it looks bigger, i.e. its perspective/ stimulus size gets bigger, than the actual physically bigger disc.

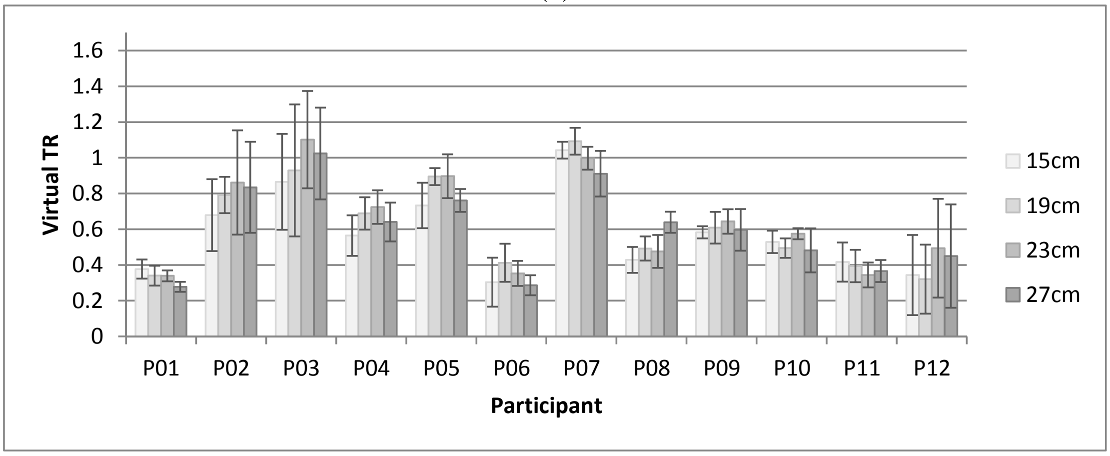
However, during the pilots it was a common reaction of the participant to look confused but agree anyway or to disagree. As mentioned before, being unsure of the intention of the instruction we were unable to clarify and did not want to influence their decision. The final instruction, *"Now I want to discover the exact setting at which this disc* ***looks*** *the same size as the other"*, asks the participant to match the perspective sizes of the discs, where it was left up to the participant to decide what 'looks' meant. Thouless meant that the participant should report perspective disc sizes but when asked to report the size an object 'looks' the instinctive reaction is to report its real size or objectively match objects i.e. attempt size constancy. This can be corroborated by the work of Joynson (1949), which states that it is actually not natural to perform perspective matching since it requires an analytical approach to break the more innate reaction to report real size. As pointed out by Gilinsky (1955), for such matching experiments it is of utmost importance to give clear instructions on whether perspective or objective matching is expected of the participant. As a direct consequence of the unclear instructions, the type of matching cannot be determined and during the experiment it became apparent that each participant may have performed the task differently. This left us with results that we are unable to interpret as we have no way to determine whether they were matching perspective or objective size.

The results of the pilot experiment in Figure 3.9 shows a change in TR on differing cue conditions goes in both directions. The TR values of participants P01, P04, P05 and P12 decrease in the virtual condition, whereas, all other participants show an increase in

TR. This is in contrast to the result that Thouless (1932) reported — manipulating the cue conditions resulted in the movement of TR values in a single direction. The two directions of movement in our experiment potentially identifies 2 groups of participant, one performing perspective matching and the other objective. However, as mentioned before this is not verifiable.



(a)



(b)

Fig. 3.9 Bar charts showing the mean TR values per participant for each disc in the real and virtual conditions, arranged in ascending order of overall mean value in the physical condition. Error bars denote $\pm\sigma_t$, the trial standard deviation for each condition tested. The labelling along the horizontal axis is the anonymised identifier of each participant.

3.2.3 Difficulty understanding Thouless' terminology

As mentioned before, a key misunderstanding was the lack of clarity of what P actually was. Part of the problem was that Thouless (1931a) did not use consistent terminology. For instance, some of the terms he uses to describe P are: *"character seen"*, *"character perceived"*, *"character of perception"*, *"phenomenal character"* and *"apparent character"*.

Initially the terms were believed to suggest the final understanding of the object itself, i.e. its real character. In other words when he talked about the phenomenal character of size we thought he was talking about size constancy. In fact, as we now know, he was talking about the perception of the size of the image subtended to the eye. This was only understood after carrying out the re-examination of Thouless' size-distance experiment as detailed in section 3.2.4.

This was an established problem as highlighted by Joynson (1949), who identified that different fields studying human vision use the same term, 'apparent size', to infer different meanings. One is angular size, that is, the visual angle (or what we have been referring to as perspective size), another is size constancy, that is, the real physical size and the final is Thouless' PR, which is a compromise between the other two.

3.2.4 Re-examination of original Thouless size-distance experiment

Following the pilot experiments it was clear that any future experiment must use clear instructions of the matching task that must be performed. However, since we were still unsure of the intentions of the task, we re-examined the original papers detailing the size-distance experiment.

A crude set up of the apparatus was made to match that of the Thouless experiment consisting of two discs of different sizes, placed at different distances, at eye-level facing the observer (Figure 3.10). Both discs are placed in the observer's field of view, making sure that they do not overlap. Since Thouless spoke of how PR was affected by stereo cues, the observer adjusts their position until the size of both discs match perspective sizes with one eye closed. When both eyes were then used to view the discs, it was found that they cease to appear to match and that the perspective sizes clearly looked different. This led to the realisation that the phenomenal character is not a final understanding of the real object's size, i.e. size constancy, but is instead a perception of the retinal stimulus. Here was a new finding. The judgement-process of size-constancy was familiar and we imagined conversely that the retinal stimuli could simply be subjected to measurement. It had not been previously contemplated that the retinal images we 'see' are themselves a result of some higher-level process. That the brain takes an object's size cue and perspective cue and comes to some weighted combination to produce the perceived end result.

Volunteers were then invited to try the procedure and they reported the same effect of the perspective sizes of the disc shrinking or growing depending on monocular or stereo viewing. Until this point it had not been contemplated that the task may have been about perspective matching, based on the assumption that a task matching the perspective size of two discs would be trivial.

It is worth mentioning that Thouless at no point ever comments on just how remarkable it is that participants cannot match the perspective sizes of two stimuli. However, we can now be certain that Thouless instructions were intended to have the participant attempt perspective matching and not objective matching as previously assumed.

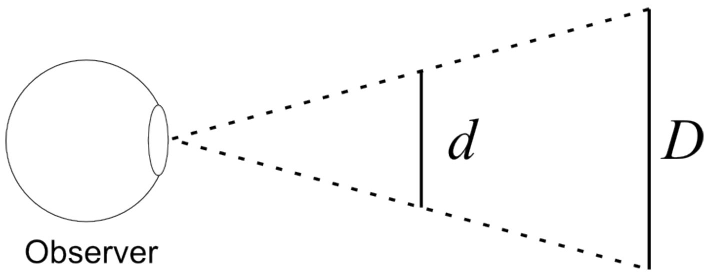


Fig. 3.10 Overview of the disc arrangement to re-examine Thouless' size-distance experiment. A 27cm white felt disc *D* was fixed to a wall and a 19cm disc *d* was fixed upright on a tripod, both discs were adjusted to be at the same height and did not occlude each other. With one eye closed the observer initially positions themselves such that discs *d* and *D* are not overlapping and appear to match in perspective size.

3.2.5 Conclusion drawn from the exploratory work

Although on reading back through the original work it is now clear what Thouless was describing, this understanding was achieved only on conducting the practical work. These pilots were also essential for the design of the experimental methodology of later experiments, aspects of which are reflected upon in later chapters and not explicitly mentioned here. Additionally, the exploratory pilots have resulted in an understanding of PR that was difficult to garner from just reading the papers. The results show first hand observation of PR, confirming that there is something to measure.

It is now ascertained that the goal of the original experiment instruction is perspective matching and that there is an effect to measure. It is further found that it does not matter if both stimuli are in the field of view. As long as objects are viewed with both eyes fully open and focused upon the discs, PR cannot be defeated.

An important outcome of the pilot experiments is the identification of the need for a better way to instruct a participant to perform perspective matching for future experiments. It was also found that it would be more practical to test both the VE and Real conditions during a single session within a MR space since it mitigates the concern surrounding the temporal stability of PR.

3.3 Summarised properties of PR

Based upon our informed understanding Thouless' work, this chapter concludes by summarising the properties of PR.

PR is subjective showing individual differences in TR. Thouless (1932, 1933) found that these differences are due to traits of individual psychology. Treating the brain as a 'blackbox' (Figure 3.6) these can be regarded as internal factors that influence PR.

PR is dependent on an awareness of the physical properties of the real object like its physical size and distance from the observer. This awareness is not due to previous knowledge, but on perceptual indicators i.e. depth cues in the case of size. Thouless (1932) demonstrated that it is responsive to changes in stereo cue conditions. Treating the brain as a 'blackbox'(Figure 3.6) these can be regarded as external factors that influence PR.

PR is an innate, involuntary and unmodifiable response to visual stimuli. PR is not mediated by experience, skill or previous training; participants cannot get better at it with practice. When measuring TR it cannot be consciously manipulated as long as the participant performs the matching task with both eyes fully open and focused upon the stimuli.

In conclusion this chapter has fulfilled objective [ii] of the research aim by identifying PR as a directly measurable property characterising subjective perceptual experience of visual stimuli.

Chapter 4

Hypothesis of the PhD

As stated in Chapter 1, the aim of the research is to identify a direct measure of VE quality that relies on participants' innate perceptual responses to visual stimuli. This chapter discusses how the background research and exploratory work addresses the objectives (as restated below), based on which the hypothesis of the PhD research is stated. This is followed by a list of presumptive key properties of TR that must be demonstrated in order to prove the hypothesis.

4.1 Research objectives restated

The first two research objectives in the introduction chapter are as follows:

i. Identify a key element of spatial perception where there is a suspected deficit in the VE compared to the real world.

ii. Identify a directly measurable property capable of characterising subjective perceptual experience of visual stimuli.

As required by the first objective, the research background in Chapter 2 identifies the deficit of depth perception in a VE and discusses the body of work pertaining to it. Furthermore, the type of VE to be studied, namely a large surface display VE, is identified.

The directly measurable property required by the second objective is identified in Chapter 3 as the psychological trait phenomenal regression, with TR as its numerical measure. PR describes how the perception of retinal stimulus of an object's properties (i.e. size, shape and brightness) is influenced by the perception of the object's real physical attributes. As detailed in Chapter 3, Thouless (1931a, 1932) investigated and developed a methodology to measure TR of perspective size and found it to rely on an awareness of the object's physical size. That awareness of the physical size is dependent on the available visual cues for forming a perception of the distance to the object. Therefore, PR of perspective size is an element of visual perception that is suspected to be affected by deficits of depth cues in a VE. The third objective of the research can now be restated as:

iii. Within the scope of depth perception for large-screen immersive display VE, investigate the performance of TR to establish its value as a VE assessment metric.

4.2 Hypothesis

Based on the aim, the fulfilled objectives ([i] & [ii]) and the requirement of the remaining objective, the hypothesis of the research can thus be stated as follows:

PR is a potentially useful measure of veridical perception to compare visual responses to virtual and physical stimuli.

In order to prove the hypothesis, key properties of TR must be established to be true. These properties are furnished by the experiments detailed in Chapter 5 and Chapter 6.

4.3 Presumptive key properties of TR

For TR to be established as a robust measure of veridical visual perception in a VE, the research must demonstrate the following key properties:

- TR is a valid measure. This involves verifying that there is something to measure and that PR occurs with both physical and virtual stimuli.
- TR is a reliable measure. This involves verifying that measuring under constant cue conditions produces consistent results. TR may or may not have long term temporal stability, however, PR must be reasonably stable for the duration of testing.
- TR is responsive to changes of cue conditions in a VE. This involves verifying that an individual's TR is affected by the depth cues changes in a VE while using the same measurement process and participants.

Chapter 5

Comparing PR between virtual and physical stimuli

This chapter describes the motivation, design and results of an experiment comparing individual TR of perspective size between virtual and physical stimuli. In order for PR to occur there must be some perception of real size of the discs being matched and therefore the distances. Given the identified deficit of distance perception in a VE, it is suspected that the results will show that TR differs between a virtual and physical disc. The experiment addresses objective [iii] of the research aim. It contributes to proving the hypothesis by attempting to verify the presumptive key properties of TR, that it is valid, reliable and responsive

The chapter is structured as follows: Section 5.1 gives an overview of the design of the experiment. Drawing on the lessons learnt from the previous pilot studies, it describes the task that participants will conduct. This is followed by an overview of the experimental factors, disc type and size in section 5.2 and a description of the VE used in this study in section 5.3.

Section 5.4 describes the arrangement of the apparatus, justifying the chosen disc distances from the participant and their position in the participant's field of view. Section 5.5 is a discussion of the generation of the standard and response discs.

Section 5.6 provides details of the experiment procedure which is split into three phases. The first is the instruction of the participants to attempt perspective size matching, which is one of the most important aspects of the experiment procedure as objective size matching would invalidate the results. The second phase is the calibration of the equipment to achieve the correct disc positions and distances. The third phase is the testing itself.

Section 5.7 presents the results, initially exploring the data visually with bar graphs and by plotting the estimated marginal means (a profile plot). Following this is an analysis of variance (ANOVA) to test the statistical significance of any differences due to the experimental factors and a study of correlation and its significance. Additionally the relationship between TR and age of the participant is examined. Section 5.8 then discusses the conclusions that have been drawn from the result and finally section 5.9 concludes with a description of the key properties of TR that are established by the experiment.

5.1 Experiment design

The experiment is designed to measure and compare TR in different individuals by using virtual and physical stimuli. The participant is tasked to adjust the diameter of a disc D at distance L to match the perspective size of disc d displayed at l (Figure 5.1). The disc that is adjusted by the participant is referred to as the 'response' disc, the disc that is used for comparison is the 'standard'. This arrangement of discs is the other way round to the Chapter 3 analysis, but as discussed a property of the TR equation is that it does not matter which disc is the response and which is the standard.

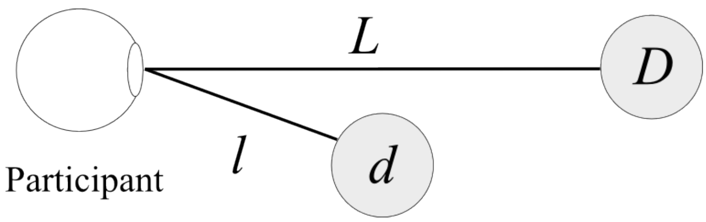


Fig. 5.1 The arrangement of the response and standard discs. The participant adjusts the size of response disc D at distance L until it matches perspective size of the standard disc d at distance l.

There are 2 matching tasks; the first is to match a physical response to the virtual standard. The second is to match a physical response to a physical standard. Both tasks will be performed in the same sitting within a single MR space as illustrated by Figure 5.2.

Testing both the physical and virtual disc conditions within a single experiment, as opposed to 2 separate experiments, resolves the issues experienced when administering the exploratory pilot experiments. It alleviates additional effort and potential difficulties of ensuring both VE and real set-up are equivalent. A single set-up is faster to build and administer. It ensures that the participant is performing the same matching task for both disc conditions. Testing in a single sitting neutralises any potential effect of time between the experiments.

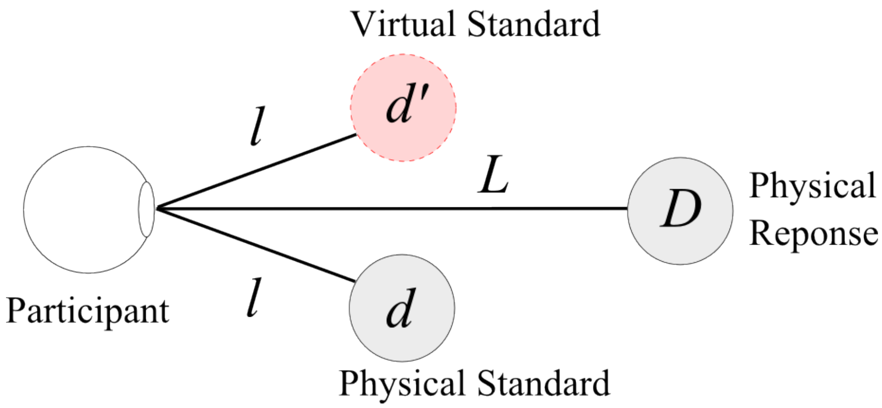


Fig. 5.2 The arrangement of the physical standard disc d, virtual standard disc d' and physical response disc D within a single experiment set-up.

5.2 Experimental factors

The largest disc that can be displayed on the apparatus was 27cm in diameter and successive disc sizes decrease by 4cm intervals. Participants are thus tested with 4 sizes of disc with diameters of 15cm, 19cm, 23cm and 27cm.

The participant is tested in both disc conditions, therefore, the experiment is described as a within-subjects design. There are 2×4 factors making a total of 8 repeated measures. Since it was estimated that a participant would take 1 minute per trial and we want to limit participant fatigue and boredom each repeated measure is sampled 5 times. This makes a total of 40 trials per participant that are selected at random during the experiment. The experiment factors are summarised in Table 6.1.

Table 5.1 The experiment has two independent variables: type, with two levels (physical, virtual) and size with four levels (15cm, 19cm, 23cm, 27cm). There is one dependant variable TR. Participants are tested in each factor making this a 2×4 within subjects design, with a total of 8 repeated measures, each of which is sampled 5 times. Making a total of 40 trials per participant, selected at random during testing.

Type	**Size**
Virtual	15cm 19cm 23cm 27cm
Physical	15cm 19cm 23cm 27cm

5.3 Description of VE used in this study

The Hull Immersive Visualization Environment (HIVE) is an optically motion-tracked room with a raised stage area, with a depth of 3.22m that is 5.33m wide, in front of an auditorium-type seating area that faces a wall-to-wall display that is 5.33m wide and 2.44m high. The single-surface display is rear-projected and driven by two horizontally placed, active-stereo projectors, with the blending of any resulting edges performed by hardware.

The user wears a pair of optically-tracked, LCD shutter glasses the position and orientation of which are used to calculate the display view. The eye positions are used to calculate the asymmetric parallel axis stereoscopic viewing frusta. In addition head orientation is taken into account, collectively generating perspective, motion parallax and binocular disparity cues to depth. As a result the depth at which the user perceives an object is at the vergence point that the pair of stereo images are fused. Vergence, in the real world, works in conjunction with accommodation of the eye's lens to focus the light rays coming from the object onto the retina. However, on a screen projection implementation of a virtual world, the light rays always emanate from the display surface which creates an accommodation-vergence mismatch for the user. This accommodation-vergence mismatch gets progressively worse with increasing distance from the screen.

5.4 Apparatus arrangement

According to the space classifications by Cutting and Vishton (1995), the experiment can be classified as taking place in action space. Action space is defined as the space beyond personal space but within 30m, where personal space is defined as the distances within and slightly beyond an arm's length. Both the standard and response discs are set-up at distances within the action space. This is in part due to the desire to keep the experiment set up as close to that of Thouless' original experiment but dictated by the physical constraints of the HIVE.

The distances of the standard discs (d and d') and the response disc (D) are nominally set up at 1.22m and 2.72m away from the participant. Given the required separation of the physical and virtual discs, this is the closest to the participant that the disc can be created without the edge of the screen causing occlusion and thus disrupting the depth illusion. The angle subtended between a standard disc and the response disc, from the participant's position during a trial, is 15°. The apparatus arrangement is illustrated in Figure 5.3.

When the task is performed correctly the participant looks at and focuses on each disc, one at a time, with both eyes. However, during the pilot trials participants reported difficulty in following the instructions, forming a strategy of placing both discs in their field of view while focusing on neither. This has the undesirable effect of eliminating the binocular cues of focus and vergence, making perspective matching easier to perform. The solution adopted to prevent such behaviour in this experiment is to render a disc only when the participant directly faces it, i.e. their view direction is perpendicular to the disc surface.

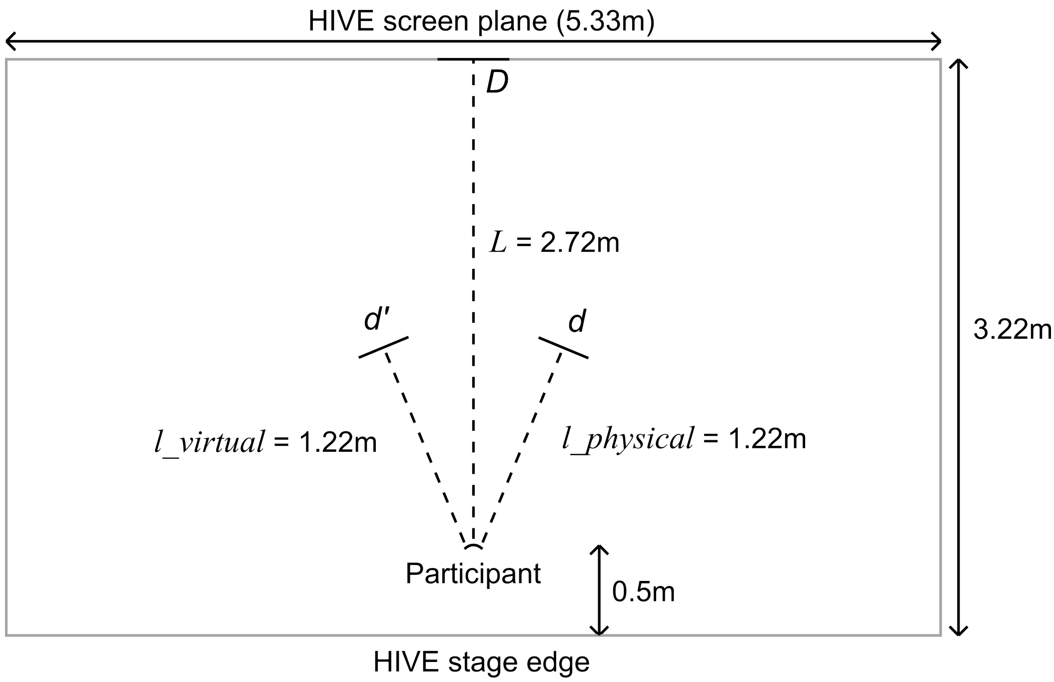


Fig. 5.3 Plan view of the experiment set-up on HIVE stage.

5.5 Generating the standard and response discs

The experiment consists of three discs projected on three displays: the HIVE screen used as a large physical monitor on which the response disc is projected, a standard LED monitor on which the real standard disc is projected and a virtual monitor implemented using the artificial depth cues on which which the virtual standard disc is projected (Figure 5.4). The monitor displays are positioned in the environment so they do not overlap in the participant's field of view and when faced by the participant are orientated at such an angle that they are perpendicular to the viewer. All three discs are white in colour with no texture. A light meter was used, through the shutter glasses, at each display to ensure uniform brightness (within a range of ± 1 lux) across the three discs.

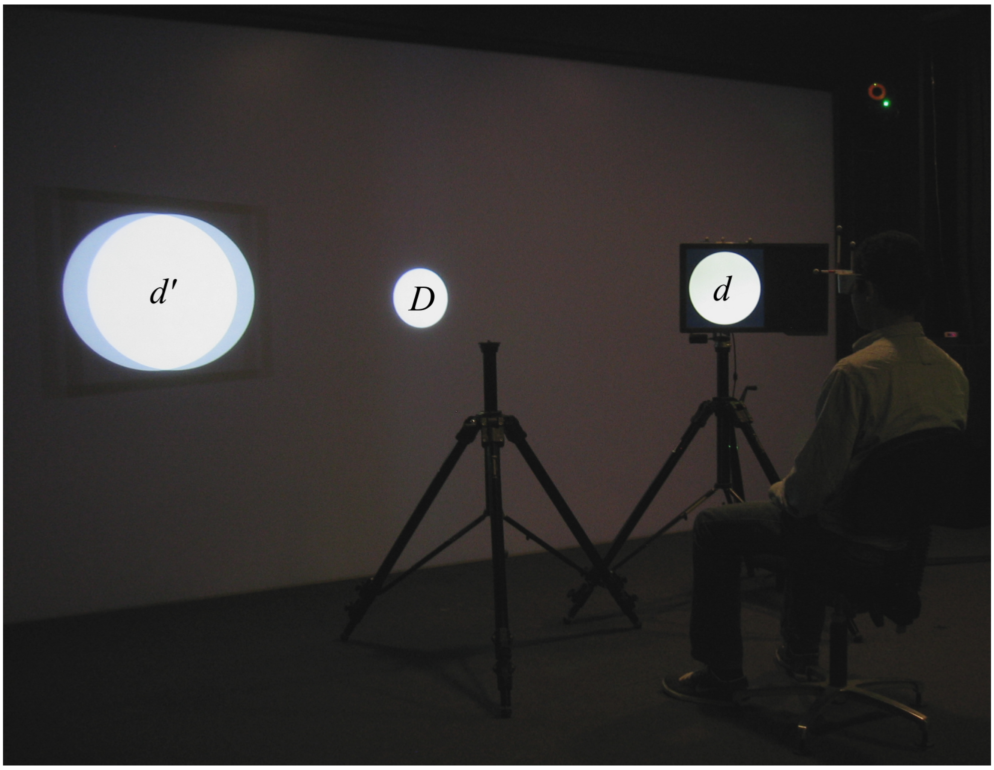


Fig. 5.4 Photograph of the experiment apparatus. The physical standard disc d is displayed on a monitor fixed to a tripod. The virtual standard d' is displayed on a virtual monitor that is perceived to be fixed upon a physical tripod. The response disc D is displayed on the HIVE screen. For illustrative purposes all three discs are displayed, during an experimental trial disc D and either disc d or d' would be visible. Note that the lighting conditions of the photograph are not illustrative of the actual experimental conditions. Aside from labelling, the photograph image has not been manipulated.

5.5.1 Physical response disc

As mentioned the response disc is adjusted by the participant to match a standard disc. It is displayed on the HIVE screen to eliminate any extraneous size cues that would be available from a conventional monitor surround. The central response disc is displayed on the plane of the HIVE screen with zero parallax and no stereo separation. The focus and vergence cues are thus correct here so that the disc has the perceptual properties of being physical. As mentioned before, when the participant faces this disc it is 2.72m from them. At this distance the there are little focal (accommodation) and vergence cues left as the eye's muscles are relaxed.

5.5.2 Physical standard disc

Like the response disc, a disc on an LED monitor has the perceptual properties of being physical on account of accurate focus and vergence cues. The LED monitor is supported on a tripod to the right of the participant at a distance of 1.22m away. Although the physical dimensions of the monitor are 56.8cm by 33cm, the screen is masked by black felt to reveal only an area of 29cm by 29cm to maintain symmetry with the virtual standard disc. The LED monitor is tracked to ensure symmetry of the experiment set-up.

5.5.3 Virtual standard disc

The virtual standard disc is displayed on a virtual monitor. The virtual monitor, a black rectangle corresponding to the visible portion of the physical LED monitor, is projected to the left of the participant mirroring the position of the LED monitor. The virtual monitor is displayed on the HIVE screen with negative parallax so that it appears to be 1.22m from the participant. A physical tripod is positioned under it such that the virtual monitor appears to sit on it (Figure 5.5). The inclusion of both the virtual monitor and the physical tripod is necessary to ensure that the virtual condition was matched to the physical.

The stereo depth cues of the virtual disc suffer from an accommodation-vergence mismatch (Figure 5.6). Upon fusing the offset left-eye and right-eye images of the virtual disc, the view direction of the participant's eyes converge at the vergence point *vp*, where *vp* is at the distance of 122cm from the participant. However, their eyes will accommodate to the focal point *fp* 297cm away at the HIVE screen.

The virtual disc suffers from crosstalk and, therefore, perceivable ghosting if viewed through the edges of glasses by using peripheral vision. This effect is exacerbated by the brightness of the disc. The ghosting effect disappears when the virtual disc is viewed straight-on directly through the lenses. The artificial blinker mechanism previously discussed in section 5.4 mitigates this ghosting by preventing the user from viewing the disc using their peripheral vision. Note that the general level of crosstalk inherent in the HIVE system was not measured.

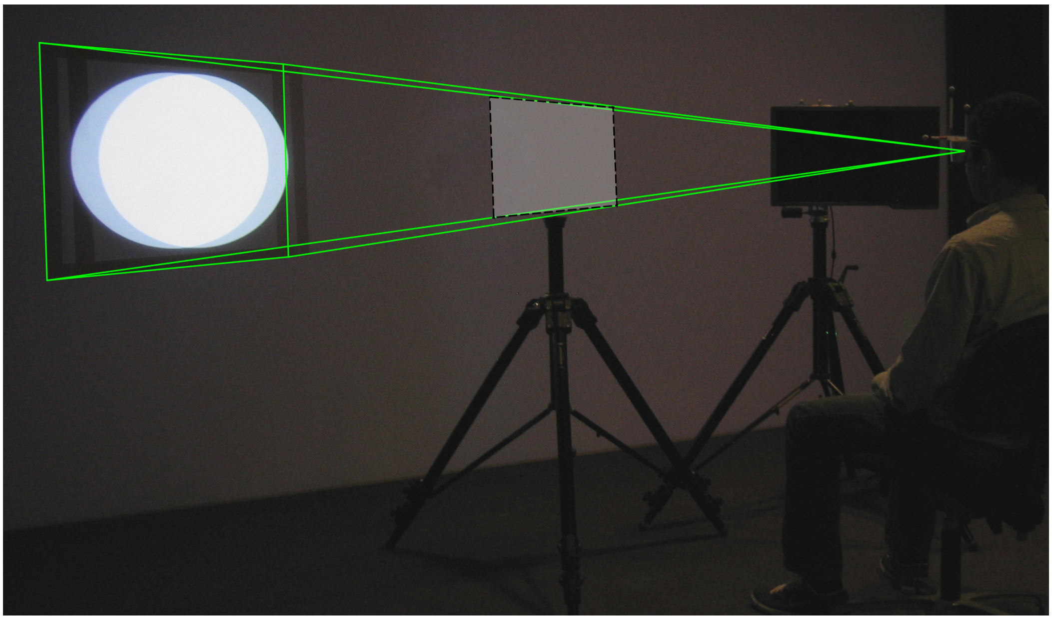


Fig. 5.5 The green lines illustrate the projection of the virtual monitor and disc d' from the HIVE screen to the participant's eyes. The combination of stereoscopy and motion parallax cues make it appear to rest upon the physical tripod, as indicated by the grey rectangle. For clarity discs d and D have been removed from the photograph and the borders of the virtual monitor have been enhanced in this image.

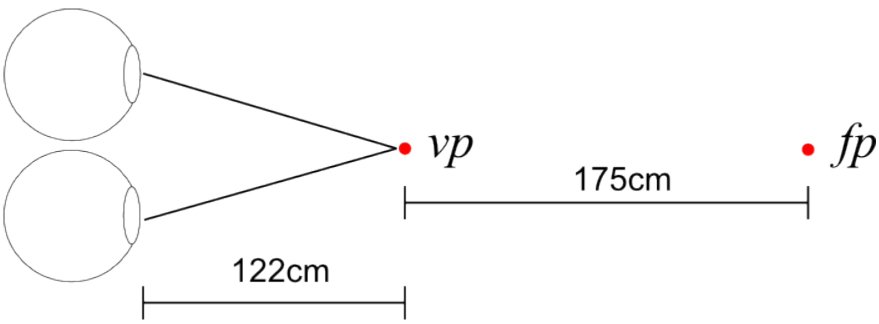


Fig. 5.6 The stereo depth cues of the virtual disc suffer from an accommodation-vergence mismatch. The eyes converge at vergence point (vp) at distance 122cm, but accommodate to the focal point (fp) at distance 297cm.

5.6 Experiment procedures

There are 3 phases to the experiment, instructions, calibration and testing, which we discuss here.

5.6.1 Instructing participant to attempt perspective matching

The goal of the task is for the participant to adjust the perspective size of the response disc to match the perspective size of the standard. Explicit instructions must be made to ensure that participants understand the task they are to perform is perspective matching, not objective. This is one of the most important aspects of the experiment; insufficient instructions in the pilots resulted in data that could not be interpreted. Participants are instructed in the same manner as Thouless (1932) but with some modifications to clarify meaning.

The participant is invited to sit on a chair on the HIVE stage. They swivel to the left of the stage and face the the black curtained wall. A 27cm white disc made from felt is fixed to to the curtain at approximately their eye level — with their verbal guidance. The experimenter then holds a 19cm disc, at the same distance next to the 27cm disc, at the participant's eye level. The experimenter states *"This is a large disc, this is a small disc, this one looks larger than the other, doesn't it?"* The point of this question is to make it clear that the discs are physically different sizes and that at the same distance they produce different retinal images. The participant acknowledges this.

The experimenter then carries the disc they are holding, the smaller 19cm disc, towards the participant, being careful to keep it at the same eye level as the larger disc and not occluding the larger disc. The experiment states *"As I bring this disc closer to your eyes it looks bigger in your field of view until finally it looks bigger than the other disc"*.

This is similar to the instruction given in the pilot experiment, but with the additional phrase *"in your field of view"*. The addition of this phrase makes it clearer to participants that the smaller disc looks bigger — it subtends at a larger visual angle to their eyes.

The experimenter then says *"Now if I slowly move the disc away from you there is a distance where both discs look the same size, i.e. where their perspective sizes match?"*. The addition of "where their perspective sizes match" improved understanding. The experimenter would move the disc away from the participant until they acknowledged a distance where the statement is true.

Finally the experimenter says *"During the experiment you will perform a similar task, but instead of moving a disc, both discs will be fixed at different distances. You will adjust the diameter of a far disc until its perspective size matches a near disc."*

5.6.2 Calibration of apparatus

In order to generate the correct stereo separation of virtual images for a participant, it must be configured to match inter-pupillary distance (IPD) of the participant. Thus prior to the experiment, a participant's (IPD) is measured, first using software that uses a web-cam and then physically verified by measuring with a ruler.

The first step in the experiment set-up is to calibrate the apparatus to ensure the discs are at the required distance, height and orientation in relation to the participant. The participant is made aware that the calibration is not part of the experiment itself. The calibration involves adjusting the distance between the participant's chair and the HIVE screen. The participant is required to sit on the chair and look at each standard disc in turn to ensure that the distance from their eyes to both the standard discs is 1.22m and that to the HIVE screen is 2.72m. This may involve a few repetitions with the participant standing up while the chair is moved, taking up 5-10 adjustments to get the desired position accuracy. A participant's head is not held at a constant position throughout the experiment but instead is continuously tracked and the actual distance of the participant's head from the disc is used for each calculation of TR. It is found that all participants performed the experiment within a maximum of 7cm from their initial calibration position although it was considerably less than 7cm for most participants. Additionally, during the measurements the participants' positions remained stable, with $|l_{physical} - l_{virtual}| < 2$cm.

Once the position of the chair is fixed the participant is allowed to adjust the height of the seat so that they are comfortable and their feet are on the floor. The display software is then used to match the height of the central response disc to the participant's tracked eye level. The LED monitor that is supported on a tripod is manually lowered or raised until the participant reports that the heights of the standard disc and response disc match. This automatically causes the position of the virtual monitor, on the left, to be symmetrical and at the same height. A second tripod is placed below the virtual monitor and its height is adjusted to match that of the first tripod. A final check is made by asking the participant *"If you glance across all three discs, they should all appear to be at the same height and level with your eyes?"* (Figure 5.7). Finally the angle of the LED monitor is adjusted until it is reported to directly face the participant and the standard disc as a result appears perfectly circular. This adjustment is reflected in the orientation of the virtual monitor as a consequence of the tracking of the LED monitor.

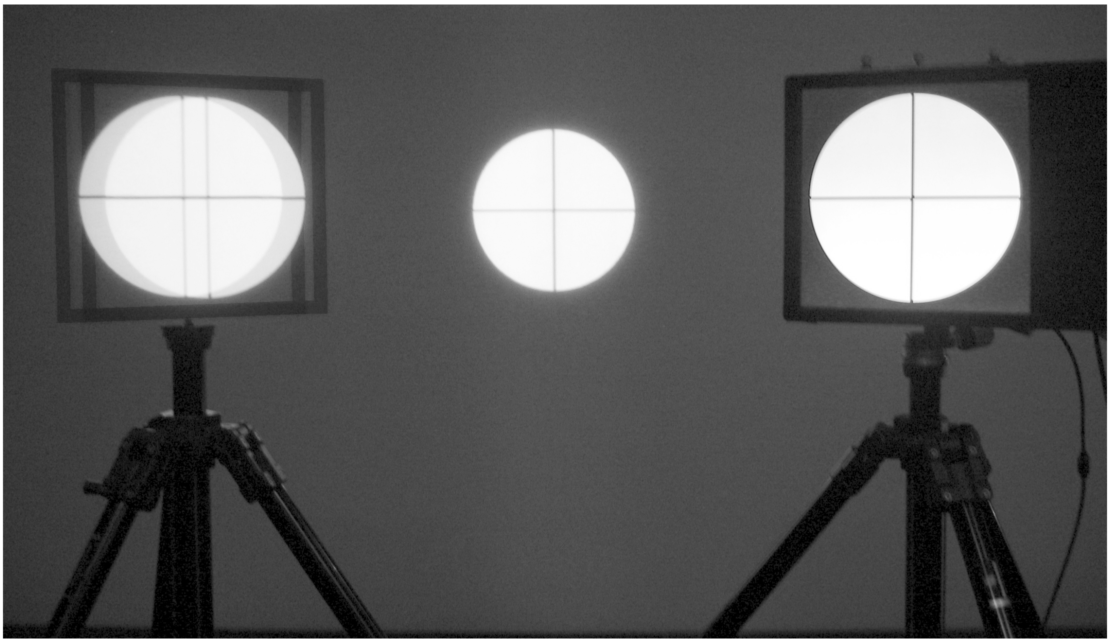


Fig. 5.7 The photograph illustrates the participant's view upon completion of the calibration procedure. All 3 discs are at the same height, face the viewer and are level with the viewer's eye. The black crosses drawn on the discs are to assist the calibration procedure. The stereo separation of the virtual disc and monitor (left-most) is not visible when the participant is wearing the LCD shutter glasses. This photograph illustrates how the virtual monitor appears to rest upon the physical tripod. The photograph has been manipulated for clarity of printing: the brightness and contrast do not match those of the experiment conditions; the image is desaturated and the borders of the virtual monitor have been enhanced.

5.6.3 Testing

The next stage of the experiment is instructing the participant on the experimental procedure. This begins with the specific instruction *“Upon each experimental trial you will be presented with two discs, one on the centre screen and one on either the left or right monitor”*. While no distinction is made between the virtual and physical monitors, it is explained that the discs are only visible when participants turn and face the display directly. The participant is instructed on the usage of the wireless mouse that is then handed to them *“With the mouse wheel you are to adjust the size of the centre disc until its perspective size matches the standard disc”*. They are then informed that all the discs will disappear between each trial and a white box with a black cross will be displayed on the central large screen for a period of three seconds, during which they are to look at the cross. They are further informed that the monitor on which the standard disc will appear will be indicated at the start of each trial (by the experimenter saying “left” or “right”) and are reminded that for every trial the adjustable response disc is displayed on the central screen.

Participants are assured that there is no limit to the number times they can turn back and forth between a standard and response disc. They are however encouraged to swivel their chair instead of turning their heads to minimise potential fatigue or strain. The participant is then informed that once they are satisfied that the response disc matches the standard disc they need to inform the experimenter who will facilitate progression to the next trial. The final instruction given is that they must not narrow or close their eyes or use any aid including their fingers for the matching task. This is followed by two practice trials, one with each monitor, during which the participant can ask questions. On completing the two practice trials the experiment begins.

5.7 Results

The participants for the experiment were chosen from postgraduate students and lecturers in the Computer Science department at the University of Hull, making sure that none work in the field or are knowledgeable of VEs, scientific visualization or psychology. However, all the participants have some exposure to head tracked stereo displays at university open day events. Sixteen participants were chosen at random — this number was determined by power analysis. The resulting group had a majority of males with just three females, the age range was 23–58 with a mean age of 39. All participants had normal or corrected-to-normal vision, however, to maintain comparability with the Thouless (1931a) experiment we deliberately did not screen for stereo vision ability.

The datasets of four participants were not included in the analysis for different reasons. The first participant's data was incomplete as the experiment was aborted due to eye strain caused by not wearing their corrective lenses. The second participant confessed to engaging in a strategy to use retinal after-images that renders the data inadmissible. The third participant forgot what they were meant to do and said they might have performed objective matches instead of perspective matches once again making the data inadmissible. The final omitted dataset was as a result of an error in the procedure calculating the wrong IPD for the experiment. Without these datasets the age range changes to 23–55.

The analysis of the results of the remaining participants, as the means and standard deviations across five trials of physical and virtual TR values per participant per disc, can be seen in Figure 5.8. The participants have been arranged in order of increasing overall mean physical TR value in order to demonstrate the range of TR values observed.

The subsections that follow analyse the data with a profile plot, tests of statistically significant effects and regression analysis.

5.7.1 Exploring the data visually with a profile plot

Plotting the mean of participant TRs of both conditions over the disc sizes provides a visual profile of the data (Figure 5.9). The plot lines do not intersect which suggests that there is no interaction effect between the disc type and disc size. The plots hints at a slight difference between the disc types, that virtual TR is a little larger than physical.

The shape of the plot lines of both the physical and virtual discs is very similar. The changes in TR over disc size appear consistent in both conditions i.e. TR increases by a similar amount between the 15cm disc and 19cm disc, and decreases by a similar amount between the 23cm disc and 27cm disc.

5.7.2 Testing for statistically significant effects

Mean TR values were subjected to a two-way repeated measures ANOVA. There were no outliers and the data was normally distributed for each group, as assessed by boxplot and Shapiro-Wilk test ($p > .05$), respectively. There was no significant interaction between disc type and disc size, $F(3, 33) = .281$, $p > .05$ and no significant main effect of disc type, $F(1, 11) = 2.852$, $p > .05$. However for the disc size factor, Mauchly's test indicated that the assumption of sphericity had been violated, χ^2 (5) = 11.874, $p = .037$, and therefore the Greenhouse-Geisser correction was applied ($\varepsilon = .669$), giving a significant main effect of

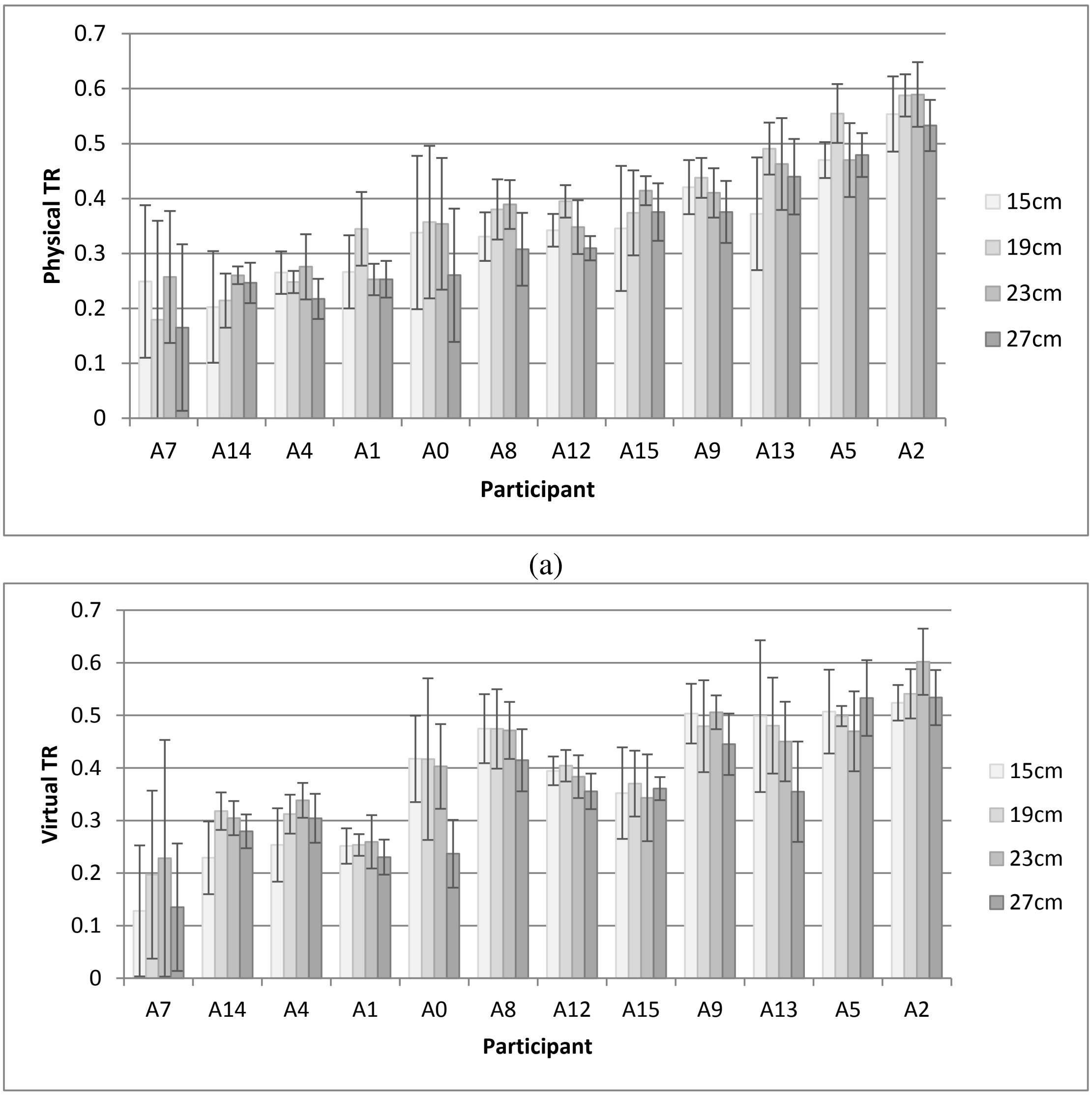


Fig. 5.8 Bar charts showing the mean TR values per participant for each disc size in the physical (a) and virtual (b) conditions, arranged in ascending order of overall mean value in the physical condition. Error bars denote $\pm\sigma_t$, the trial standard deviation for each condition tested. The labelling along the horizontal axis is the anonymised identifier of each participant.

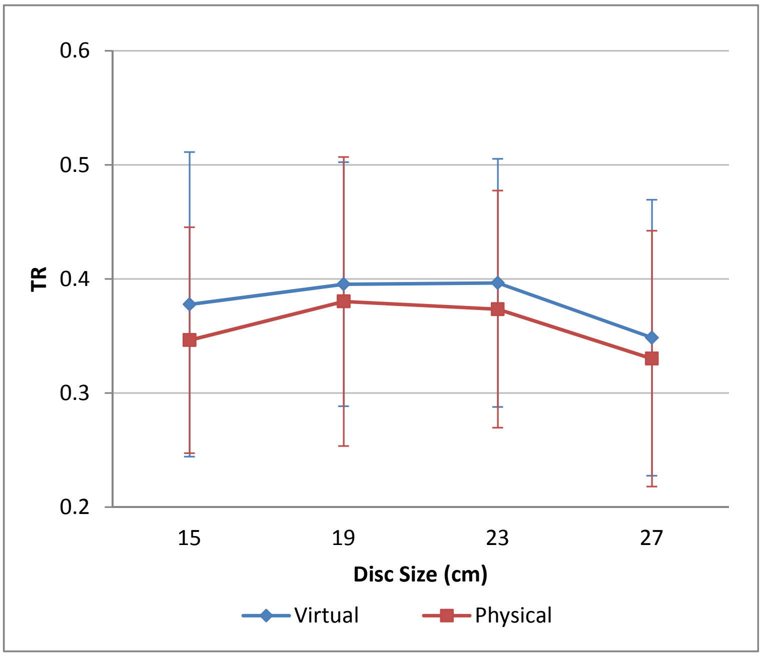


Fig. 5.9 Plots of the estimated marginal means of both the physical and virtual disc type. The error bars denote $\pm\sigma_t$, the standard deviations.

disc size, $F(2.008, 22.086) = 9.758$, $p < .05$. For the disc size, *post-hoc* pairwise analysis of TR with a Bonferroni adjustment revealed that TR increased statistically significantly from 15cm to 19cm (M = .026, 95% CL[.009, .042], p = .002). TR decreased from size 19cm to 27cm (M = .048, 95% CL[.013, .084], p = .007), and from size 23cm to 27cm (M = .046, 95% CL[.085, .006], p = .021).

5.7.3 Correlation of TR between virtual and physical discs

The correlation of TR between the virtual and physical discs are investigated by fitting a straight line. To take account of the presence of errors in both sets of measurements the approach of (Press et al., 1992, p.666–668) is adopted. As there is a significant main effect of disc size each is plotted separately; the results are displayed in Table 5.2 and the corresponding plots in Figure 5.10.

5.7.4 Examining relationship between TR and age

The correlation between TR and age is investigated by fitting a straight line. As there is a significant main effect of disc size each is plotted separately; the results are displayed in Tables 5.3 and 5.4 and the corresponding plots are in Figure 5.11.

Table 5.2 Slope, intercept, coefficient of determination (R^2) and p-value of virtual versus physical TR values, for each disc size.

Disc size	Slope	Intercept	R^2	p
15cm	0.8237	0.0295	0.7277	$< .01$
19cm	1.3108	-0.1324	0.7800	$< .01$
23cm	0.9815	-0.0135	0.8359	$< .01$
27cm	0.9804	-0.0201	0.7906	$< .01$

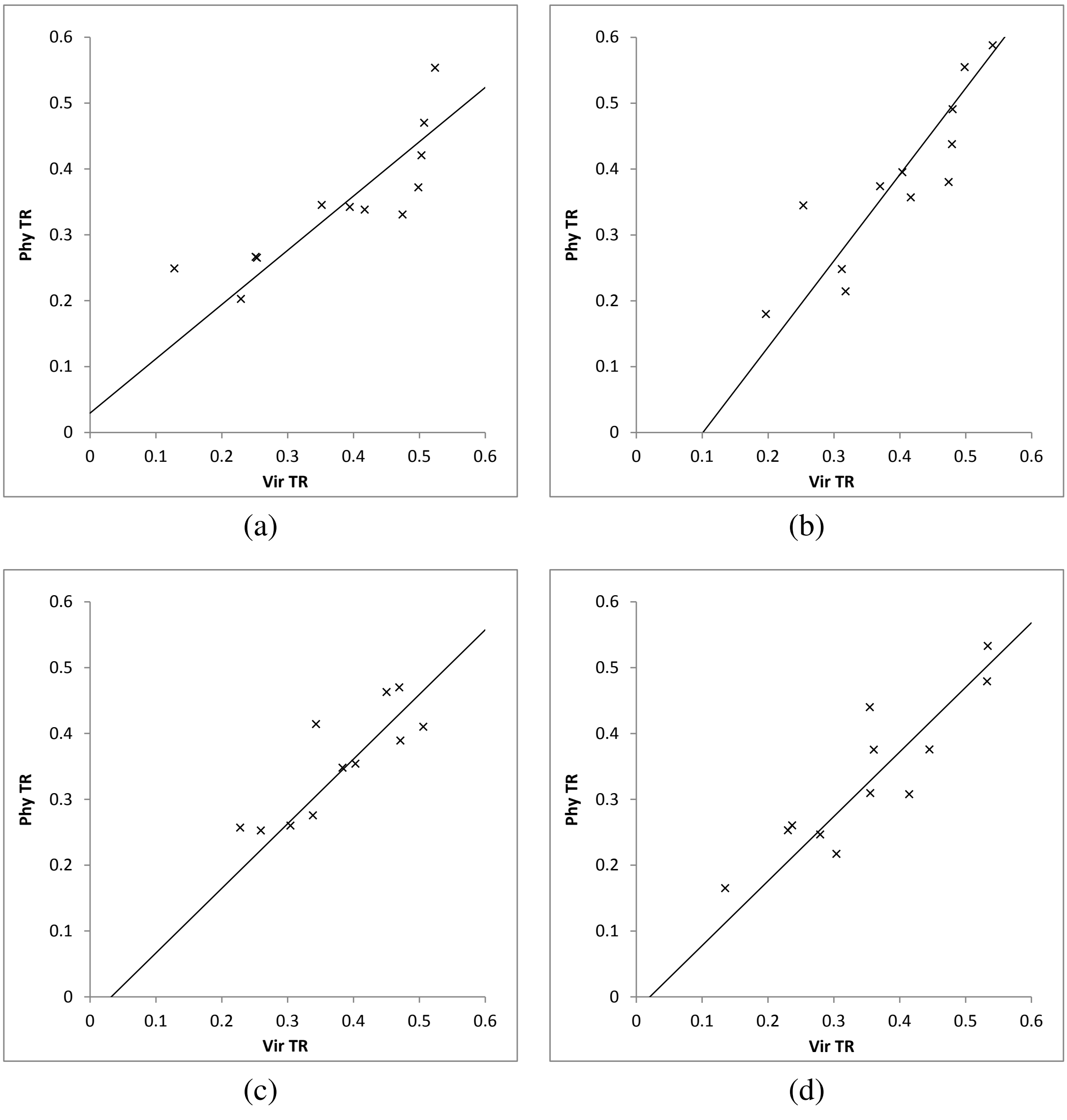


Fig. 5.10 Correlation of physical and virtual TR values for disc sizes (a) 15cm, (b) 19cm, (c) 23cm and (d) 27cm. Model-fitting parameters, weighted R^2 and p values can be found in Table 5.2.

Table 5.3 Slope, intercept, coefficient of determination (R^2) and p-value of **physical** TR values versus participant age, for each disc size.

Disc size	Slope	Intercept	R^2	p
15cm	0.0002	0.3380	0.0004	$> .10$
19cm	0.0015	0.3234	0.0125	$> .10$
23cm	-0.0001	0.3781	0.0001	$> .10$
27cm	0.0009	0.2940	0.0065	$> .10$

Table 5.4 Slope, intercept, coefficient of determination (R^2) and p-value of **virtual** TR values versus participant age, for each disc size.

Disc size	Slope	Intercept	R^2	p
15cm	0.0045	0.2028	0.1062	$> .10$
19cm	0.0040	0.2421	0.1271	$> .10$
23cm	0.0030	0.2812	0.0698	$> .10$
27cm	0.0033	0.2214	0.0684	$> .10$

Table 5.5 Slope, intercept, coefficient of determination (R^2) and p-value of **physical** TR values versus participant age, for each disc size. This data is a subset of the results within the participant age range of **23–43**.

Disc size	Slope	Intercept	R^2	p
15cm	0.0078	0.0978	0.2157	$> .10$
19cm	0.0145	-0.0892	0.4265	$< .05$
23cm	-0.0080	0.1234	0.1806	$> .10$
27cm	0.0107	-0.0110	0.2728	$> .10$

Table 5.6 Slope, intercept, coefficient of determination (R^2) and p-value of **virtual** TR values versus participant age, for each disc size. This data is a subset of the results within the participant age range of **23–43**.

Disc size	Slope	Intercept	R^2	p
15cm	0.0174	-0.2079	0.5627	$< .05$
19cm	0.0136	-0.0644	0.4808	$< .05$
23cm	0.0115	0.0129	0.3478	$< .05$
27cm	0.0153	-0.1539	0.5110	$< .05$

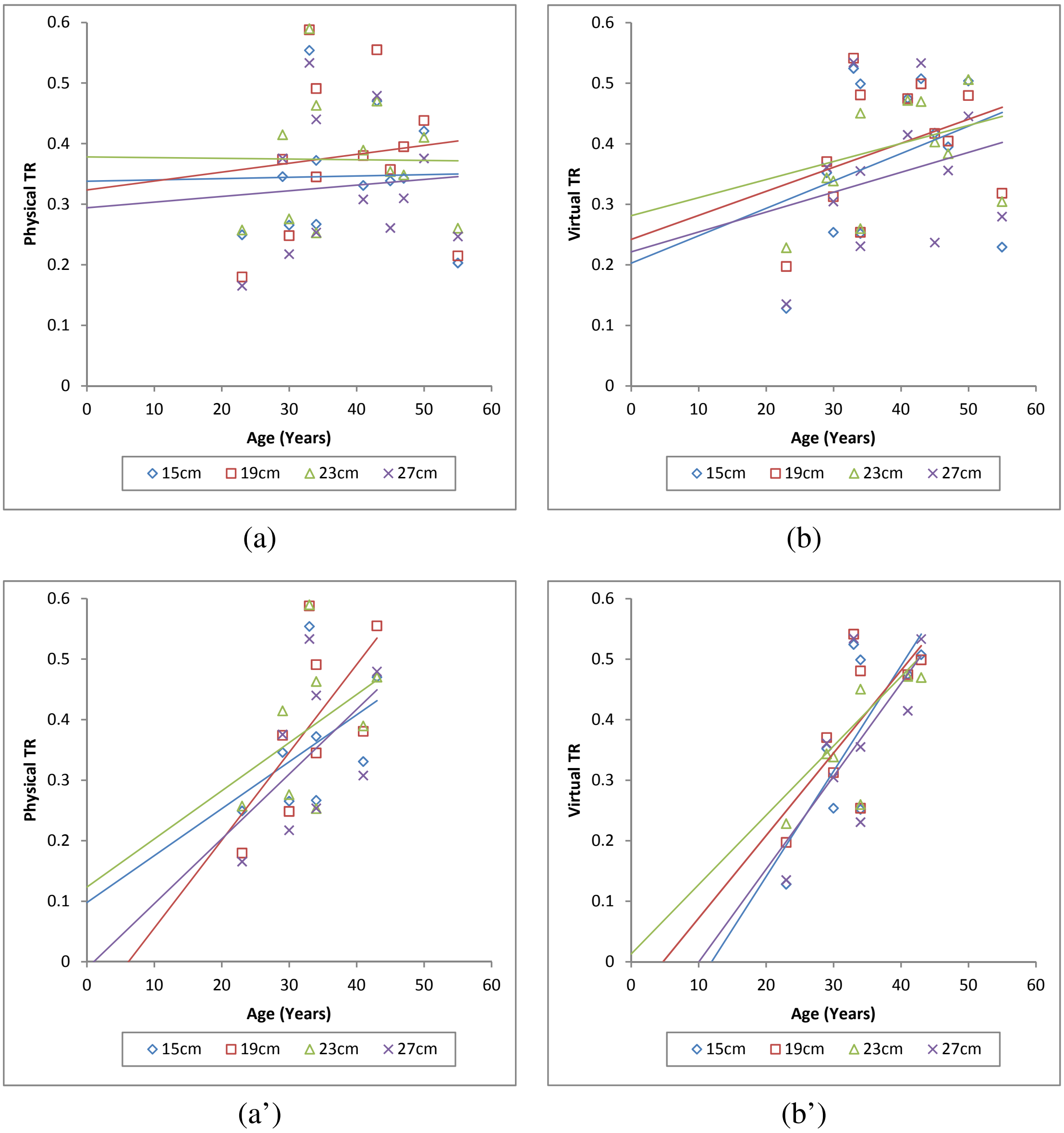


Fig. 5.11 Correlation of TR values and age for each size of (a) Physical and (b) Virtual disc. Plots (a') and (b') use a subset of the same data, but 4 participants with ages > 44 are omitted. Model-fitting parameters, weighted R^2 and p values can be found in Tables 5.3, 5.4, 5.5 and 5.6.

5.8 Discussion

The ANOVA results show no main effect on TR of disc type but do show an effect of disc size. Each of these factors are discussed in turn.

Figure 5.8 (a) shows that participants matching perspective size in the physical world exhibit $TR > 0$ as described in (Thouless, 1931a,b) and that the measure shows individual variation. In some individuals the departure from $TR = 0$ is very marked (two with mean TR around or above 0.5), bearing in mind that $TR = 1$ equates to size-constancy or perfect objective matching. We are confident this is a genuine finding and not the result of inadvertent objective matching by participants, given the precautions described in section 5.6.1 and the disclosure of a participant who realised that they had strayed from the required task. There was also a participant who could not believe they had not performed a perfect perspective match and asked to view their results after the experiment to confirm that they had not. Thouless (1931b) mentioned that some of his participants made a similar comment.

The error bars in Figure 5.8 (a) and (b) demonstrate that TR is generally reproducible for both the physical and virtual stimuli, for at least the duration of the experiment. This is true for most participants except for two individuals A0 and A7 who exhibit rather large trial standard deviations, the reasons for which require further investigation. However, indications from the other participants are encouraging, since searching for VE perceptual effects using TR require it to be stable under experimental conditions. The correlation demonstrated in Figure 5.10 between physical and virtual TR values is also encouraging since no correlation would have ruled out TR as a potential quality metric. Although the slopes and intercepts in Figure 5.10 are not precisely unity and zero these plots show sufficient relationship to warrant further investigation. Another source of encouragement is the large coefficients of determination reported in Table 5.2.

The profile plot (Figure 5.9) indicates no difference between the physical and real disc conditions, but the shape of the plots hints at a small effect of disc size. The large error bars are due to the range of individual TR values.

The results of ANOVA found no main effect on TR of disc type but found an effect of disc size. The former is unexpected but on reflection perhaps not surprising in the set-up. Given that the virtual disc is augmented by a physical tripod and we can speculate that participants accommodate to the correct distance of the virtual stimulus, even in the presence of the incorrect focus cue. Informal feedback from most participants that the virtual monitor appeared uncannily real to them supports this view. Some had not even realised it was virtual and expressed their shock when, upon completion of the experiment, they removed the shutter glasses and saw that there was actually no monitor on the left of the stage.

The latter effect of disc size is surprising and requires more investigation. This may be due to weak control of the ambient light conditions. The room lights are always set the same but the variable light wash from the HIVE screen as participants adjust the centre disc to phenomenal equality contributes to the ambient light. This is especially true for the participants with small TR, causing the far disc to be physically large even though it subtends much the same angle at the eye as the small disc.

While it was not the intention of the experiment, the relationship between TR and participant age was examined to observe any hint of a correlation between participant age and PR (Figure 5.11 (a) and (b)). The results in Tables 5.3 and 5.4 reveal no such correlation. However, due to Presbyopia, an individual's ability to accommodate quickly deteriorates from 40 years old and post 55 it is lost. Taking this into account the results were re-analysed, but 4 participants with ages > 44 were omitted leaving a subset of ages 23–43; see plots (Figures 5.11 (a') and (b')). Overall the results (Tables 5.5 and 5.6) indicates that up until $\sim$45 PR does appear to increase with participant age for both physical and virtual conditions.

5.9 Key properties of TR established

As mentioned previously, key properties of TR must be established in order to prove the hypothesis, these are: validity, reliability and responsiveness.

As evident from the experiment results PR is measurable and comparable between a physical and virtual stimuli in the specific case of the experiment. TR appears to be a valid metric, having measured the effect of PR and verifying that it is subjective, that is, it varies between individuals.

The correlated result indicates that it is reliable for the duration of an experiment.

However, for PR to be a truly useful tool it needs to be established as responsive to depth cue approximations in the VE. This was proved in the physical domain with Thouless' cue-reduction experiments. The next experiment proposes to investigate this in the VE and it remains to be seen if subtle effects, such as accommodation-vergence mismatch, have a measurable effect.

In conclusion this chapter partially fulfils objective [iii] of the research aim by investigating PR as a VE assessment metric.

Chapter 6

Effect of manipulating focal depth of virtual stimuli on PR

The results of the experiment described in the previous chapter reveal a range of TR values that, for each individual, correlates between the virtual and physical stimuli. The correlation was welcome as it establishes that, for individual users, TR can be used to compare the visual conditions of a VE to that of the real world. The values of the slope (≈ 1) and intercept (≈ 0) are a surprise given the focal cue deficit present in the virtual stimulus. It is speculated that this deficit may have been mitigated by the presence of a physical tripod placed underneath the virtual stimulus.

For PR to be a useful comparator between VE and the physical world, further experimental work is necessary to show that it is affected by changes in visual conditions in a VE. The work by Thouless (1932) in the physical realm reveals that an individual's TR is sensitive to reduction in stereo cues. In contrast this work will investigate the effect on an individuals TR values of accommodation, specifically by mitigation of the accommodation-vergence mismatch that is typically present in VEs. It is worth noting that accommodation is a stereo cue that Thouless could not manipulate, given the apparatus available to him at the time. This test of responsiveness to changes in visual conditions in a VE continues to address objective [iii] of the research aim. It contributes to proving the hypothesis by attempting to verify the presumptive key property of responsiveness of TR.

The rest of the chapter is structured as follows: section 6.1 describes an experiment that will compare TR of size between two conditions, the first is where the accommodation-vergence mismatch is mitigated and the second where it is unmitigated.

Following on, section 6.2 presents details of the apparatus arrangement for the mitigated and unmitigated mismatch conditions. Section 6.3 describes the splitting of the experiment into testing blocks that are counter-balanced.

The next sections discuss the additional measures that are taken to improve the procedure since the last experiment. Section 6.4 talks about a decision for participants to conduct the experiment unsupervised and allow them to control the progression of it. Section 6.5 then details steps taken prior to testing to verify participants' understanding of the terms objective and perspective matching.

Section 6.6 details the instruction and calibration phases of the experiment. The results of the experiment are presented in section 6.7, initially exploring the data visually with bar graphs and profile plots. Following this is an analysis of variance (ANOVA) to test the statistical significance of any differences due to the experimental factors. Additionally the relationship between TR and age of the participant is examined. Section 6.8 then discusses the conclusions that have been drawn from the result and finally section 6.9 concludes with a description of the key properties of TR that are established by the experiment.

6.1 Experiment Design

For this experiment the participant is tasked to adjust the diameter of a 'response' disc d at distance l until its perspective size matches the size of the 'standard' disc D displayed at L (Figure 6.1). This is similar to the previous experiment except the response is now the disc that is closer to the participant. As previously discussed a property of the TR equation is that it does not matter which disc is the response and which is the standard.

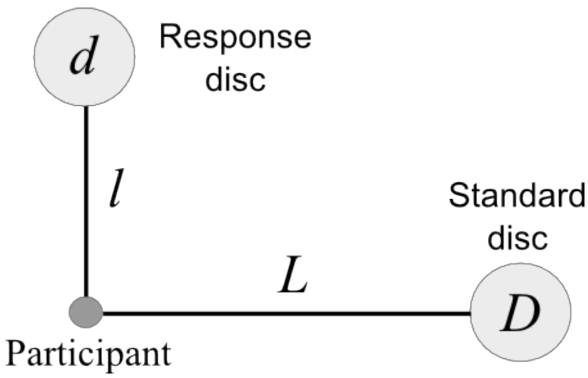


Fig. 6.1 The arrangement of the response and standard discs. The participant adjusts the size of response disc d at distance l until it matches the perspective size of the standard disc D at distance L.

Virtual objects in a VE have incorrect focal depth which creates a mismatch between the focus and vergence cues, known as accommodation-vergence mismatch. The accommodation-vergence mismatch is increasingly mitigated for virtual objects closer to the screen plane, with no mismatch for an object located at the screen plane.

The aim of this experiment is to test the responsiveness of PR to mitigation of accommodation-vergence mismatch. This is done by placing the response disc at different distances from the screen plane creating two different mismatch conditions. The first focal condition created is where the response disc *d* is placed at the screen plane. This results in the response disc appearing with the correct focal depth and no cue mismatch (Figure 6.2 (a)). For this reason, the first focal condition is called the mitigated distance condition. The second focal condition is where the response disc is created at a distance away from the screen such that the participant is at the furthest trackable distance from the screen. This results in the response disc appearing with incorrect focal depth leading to a cue mismatch (Figure 6.2 (b)). The second focal condition is thus called the unmitigated distance condition.

Table 6.1 summaries the experimental factors.

Table 6.1 The experiment has two independent variables: mismatch, with two levels (mitigated, unmitigated) and size with four levels (15cm, 19cm, 23cm, 27cm). There is one dependant variable TR. Participants are tested in each factor making this a 2×4 within subjects design, with a total of 8 repeated measures, each of which is sampled 5 times. Making a total of 40 trials per participant. The levels of the mismatch factor cannot conveniently be randomised due to the requirement to physically move equipment. Therefore they are counterbalanced. The levels of the disc size factor are randomised.

Mismatch	**Size**
Mitigated	15cm 19cm 23cm 27cm
Unmitigated	15cm 19cm 23cm 27cm

6.2 Apparatus arrangement

The VE apparatus used in this experiment is identical to that used in the previous experiment i.e the HIVE. To summarise it is a single display surface with dimensions 533cm $\times$ 244cm. The display is driven by two rear active stereo projectors, with hardware blending at the seam. LCD shutter glasses are synchronised to the left-right images displayed to

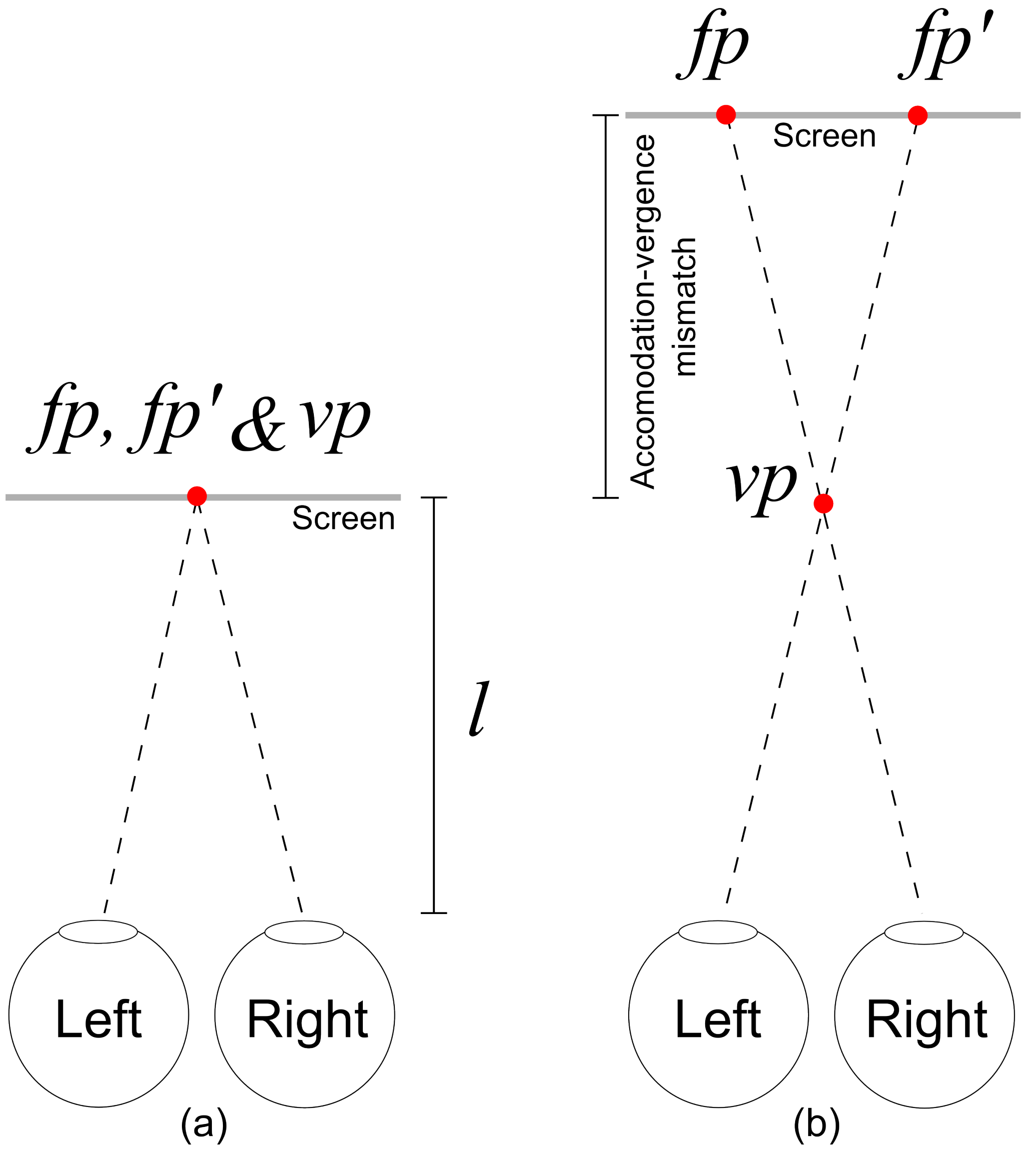


Fig. 6.2 (a) The accommodation-vergence mismatch is mitigated when a disc is drawn at the screen. The focal points of each eye (fp and fp') and the vergence point vp are at the same focal distance. (b) The accommodation-vergence mismatch is unmitigated. The eyes converge to the disc drawn at vergence point vp. However, the eyes are accommodating to focal points fp and fp'.

while optical tracking of the glasses provides the user's eye position and head orientation. Collectively these are used to generate stereoscopy, perspective and motional parallax cues. The available tracking volume encompasses the space area in front of the screen 533cm $\times$ 322cm.

The apparatus is arranged on the HIVE stage as indicated by Figure 6.3. The participant adjusts the response disc d at distance l until it matches the standard disc D at distance L. Disc d is generated by the HIVE screen and disc D is drawn on a LED monitor that is supported by a tripod.

Unlike the previous experiment a physical tripod is not placed under the disc generated by the HIVE screen, as it is speculated that it may affect the user's accommodation depth, therefore, interfere with the result. The lack of symmetry of the apparatus set-up for the two discs, with and without a supporting tripod, is not of consequence as the lighting conditions of the experiment are such that the tripod under the LED monitor is not visible.

The angle between the discs is 90°, a larger angle than in the previous experiment. This arrangement is chosen for ease of set-up, arrangement and correct calibration of the distances of D and d at each position on the stage. To perform the matching task correctly, participants must focus on each disc in turn. Despite the 90° angle it would be possible for the participants to adopt the strategy of positioning their gaze at 45° to have both discs in their field of view while focusing on neither. This is prevented using the same strategy as in the previous experiment, where a disc is displayed only when the participant is directly facing it. This is especially important in this experiment since not doing so will eliminate the stereo cues. Displaying the disc only when the participant is directly facing it also eliminates crosstalk that is perceivable if the disc is viewed through the edges of the lenses when using peripheral vision.

There was initial concern that the large 90° angle would cause participants to develop neck muscle fatigue from turning back and forth between the response and standard disc. However, during piloting it became evident that it would not be a problem as participants naturally adopt a behaviour of maintaining a fixed head position, swivelling the chair to face a disc.

An additional effort was made to control the variable ambient light levels resulting from the alteration of the response disc on the HIVE screen. This was found to be adequately controlled by the simple expedient of using grey rather than white discs.

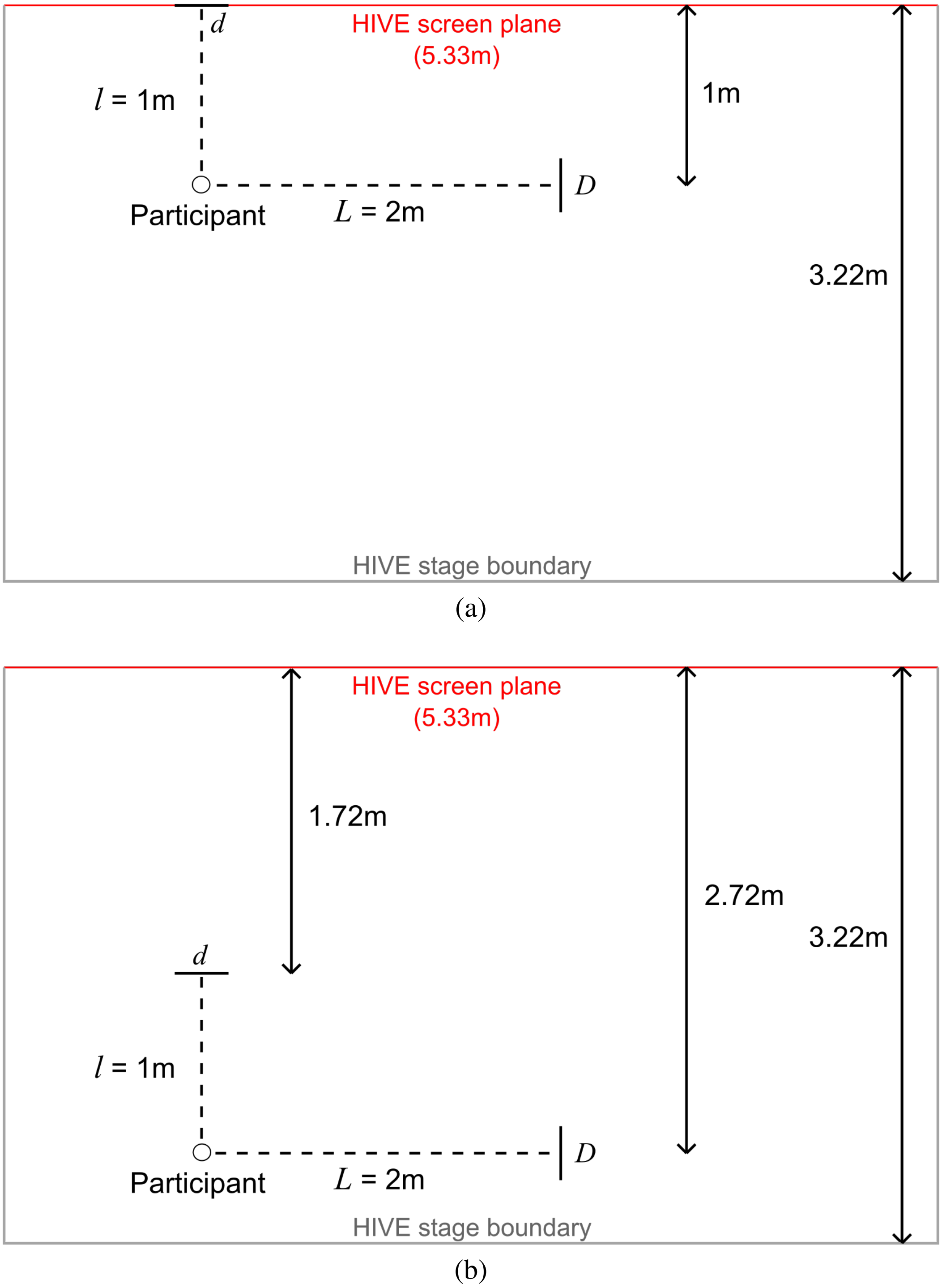


Fig. 6.3 Plan view of the experiment set-up on the HIVE stage for the (a) mitigated and (b) unmitigated mismatch conditions.

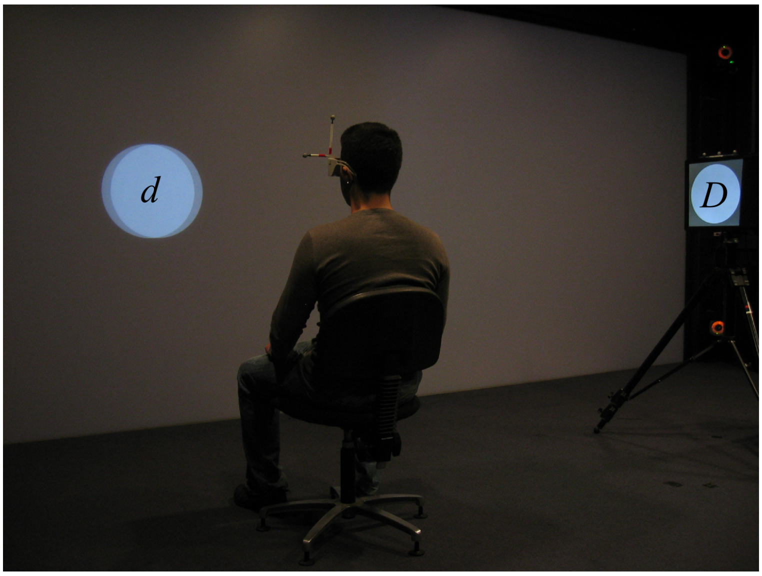


Fig. 6.4 Photograph of the experiment apparatus. The standard disc *D* is displayed on a LED monitor fixed to a tripod. The response *d* is displayed on the HIVE screen. When wearing the LCD shutter glasses the stereo separation is not visible to the participant and the disc is perceived to be 1m away. Note that the lighting conditions of the photograph are not illustrative of the actual experimental conditions. Aside from labelling the discs, the photograph image has not been manipulated.

6.3 Trial blocks — mitigated, unmitigated and partially mitigated

As there are two mismatch conditions that involve the participant being seated in two different locations on the stage, both conditions cannot be tested at once. There is always the possibility of an effect of order of trial blocks which is neutralised by counterbalancing the experiment. This simply means that half of the participants will perform the task at the mitigated distance first then the unmitigated distance and the other half perform the unmitigated distance and then the mitigated distance. Table 6.2 displays for each mismatch condition the distances to the vergence point and focal depth. Participants will perform all the trials of the task at one position, take a break, then repeat at the other.

To reduce the variability in his results, Thouless (1932) would, prior to starting the experiment, continually test the participant's PR until the TR values they exhibited were stable. In this experiment the variability in the results is reduced by having an additional block of practice trials. The participant will not be aware of the initial block being used as a variability-reduction block, the results of which will not be analysed. This additional block is called the partially mitigated distance and is positioned midway between the two extremes we are interested in testing, i.e. the mitigated distance and unmitigated distance blocks.

There are, therefore, three blocks of trials for each participant with all the participants performing the partially mitigated distance position first, with the other two blocks counter-balanced as described above. The disc size factor is randomised within each block.

Table 6.2 The accommodation-vergence values for each mismatch condition of the experiment. Showing the distance to the vergence point (VP), distance to the focal point (FP) and the distance mismatch between these two values.

Position	Distance to VP	Distance to FP	Mismatch
Mitigated	1m	1m	0
Partially mitigated	1m	1.86	0.86m
Unmitigated	1m	2.72m	1.72m

6.4 Participants are not observed during the testing procedure

During the previous experiment the experimenter was present throughout the entire testing procedure sitting at a desk behind the participant to administer the trials. Participants verbally acknowledged when they completed a trial and were ready to proceed to the next trial.

The software driving this experiment is designed to allow the participant to control when to proceed to the next trial themselves, meaning that the experimenter does not need to be present for the entire duration of the experiment. While piloting this method it was found that participants completed the task more quickly if the experimenter was not present in the room during the experiment. The results also appear to be stable suggesting that participants take the same care without rushing through the task. Participants revealed that they felt more relaxed and at ease, less like they were being watched, scrutinised and tested. We can conclude that participants feel more at ease and, therefore, more likely to give an immediate and innate response.

However, there is a risk that the lack of observation of the participant will result in less accurate performance of the task. In addition, it does mean that any queries relating to the procedure cannot be answered. The solution was to introduce an additional set of practice trials at the start of each block of trials. This practice set consisted of four trials, one for each disc size, ordered at random. These four trials are not analysed in the results, but the participant is not made aware this. The participant is observed during the the initial four practice trials and any queries are answered. At the start of the fifth trial the experimenter leaves the room for the remainder of that block of trials.

6.5 Additonal clarification of instructions for participant

Additional efforts were made to describe terms the objective and perspective matching for this experiment. Before the start of the experiment participants are greeted and invited to take a seat in the untracked seating area of the HIVE auditorium. It is made clear that the following questions are not part of the experiment, that they are not being tested, that it is simply to clarify some terminology used later during the instructions. As these are not instructions the experimenter does not follow a script. The experimenter points at a piece of the experiment apparatus, for example the PC case, and asks the participant to describe its size using their hands. The participants respond to this request in one of three ways. Some report the perspective size using their thumb and forefinger, following which the experimenter explains that they have reported the perspective size and asks them to now estimate and report the actual size. Others report its actual size by placing their hands ≈50cm apart, following which the experimenter explains that they have reported 'objective' size and ask them to report the size it looked. The third group look confused and ask the experimenter to clarify whether they are to report its actual size or how big it looks. In this case they are asked to do both and the experimenter clarifies which is a perspective match and which is an objective match. To ensure that the participant understands, the experimenter then explains the terms objective and perspective matching in the same manner discussed in the previous experiment. When compared to the previous experiment this proves to be an easier method for the experimenter to explain the terms and for the participants to understand what is required of them more quickly and with with less confusion, providing increased confidence that the terms are understood.

Thouless (1932), after instructing the participant, verified that that instructions were understood by carrying out some preliminary trials and calculating the TR value. A similar verification was considered for this experiment, but ultimately the decision was made to trust that the addition of some preliminary clarification of terms was enough.

6.6 Experiment procedure

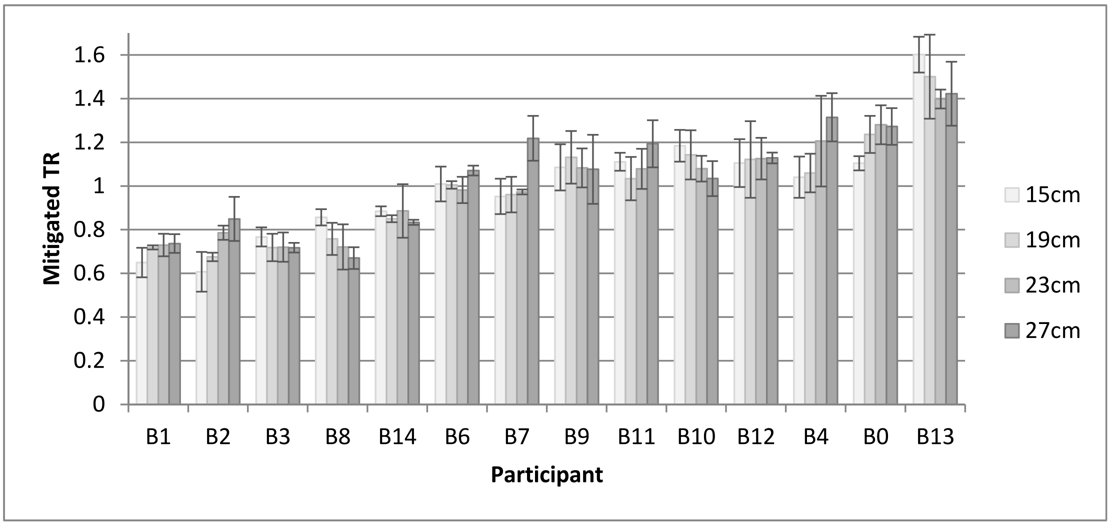
Participants are invited to take a seat on the HIVE stage to be formally instructed in the same manner and wording as the previous experiment, using the physical felt discs, as detailed in Chapter 5 i.e. *"this is a large disc, this is a small disc..."*. The final instruction is changed to *"You will adjust the diameter of a near disc until its perspective size matches a far disc. There will be 3 blocks of trials, each at a different seating position on the stage. You will take short break between each block while apparatus are adjusted"*.

Following the instruction of the participant is a calibration of the chair and LED monitor positions to correctly set the distance of the participant from the discs, L and l, and to match the disc heights to the participant's eye level. In order to do this accurately, the participant is asked to put on the tracked shutter glasses so that the distances can be measured in real-time. The participant looks at each disc in turn while the position of the chair is modified. The position of the tripod supporting the monitor is adjusted until the disc is at eye-height and directly facing them. They are asked to turn back and forth a few times while the distances are checked. It is found to be fairly quick to get the desired distances using this procedure. Nonetheless the participants head is continually tracked for the duration of the experiment, and the actual distance between the participant's eye and the disc is used for each calculation of TR.

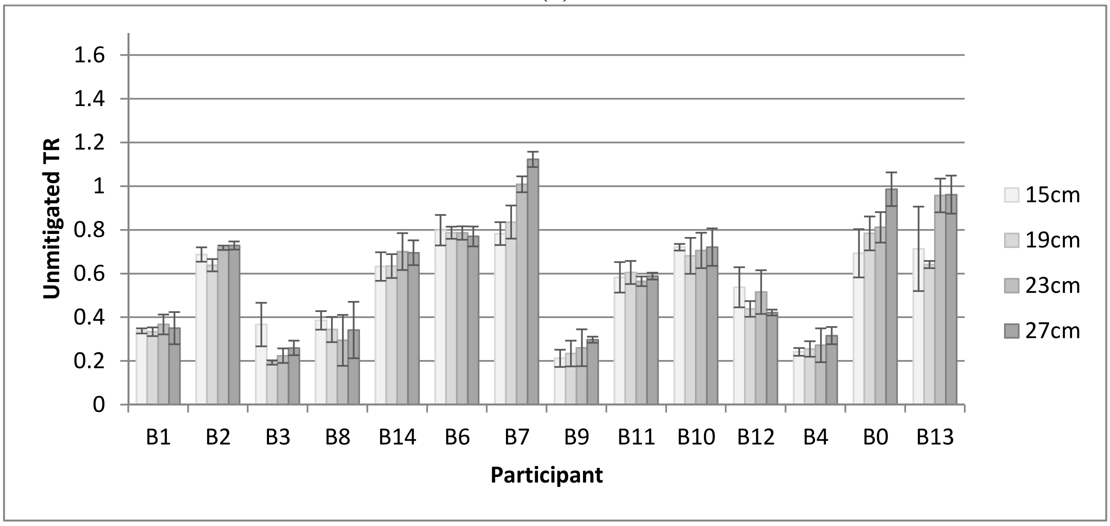
After the participant has performed the first four trials the experimenter leaves the room. Upon completion of each block of trials the participant leaves the auditorium to inform the waiting experimenter and is invited to stretch their legs or sit while the the apparatus set-up is adjusted for the next block of trials. Calibration is carried out afresh and the next block begins.

6.7 Results

The previous experiment was conducted more than 10 months prior to this experiment. In this experiment, 14 participants are drawn from the same participant population as the previous experiment, with 8 having taken part in the previous experiment. The number of participants was determined by power analysis. The participant ages range between 23 to 56 with an average of 37. All the participants performed the experiment as instructed so there was no need to omit any of the results. This may perhaps be attributed to the improvement in describing the terms objective and perspective match to the participant. The results are displayed in Figure 6.5.



(a)



(b)

Fig. 6.5 Bar charts showing the mean TR values per participant for each disc in the (a) mitigated and (b) unmitigated conditions, arranged in ascending order of overall mean value in the mitigated condition. Error bars denote $\pm\sigma_t$, the trial standard deviation for each condition tested. The labelling along the horizontal axis is the anonymised identifier of each participant.

6.7.1 Exploring the data visually

Plotting the mean of participant TRs of mismatch conditions over disc size, seen in Figure 6.6, provides a visual profile of the data. Though the results of the partially mitigated factor are not part of the analysis they are included for interest. The plot lines do not intersect which suggests that there is no interaction effect between disc size and mismatch factors.

The plot also shows that there is a large difference of TR between the mitigated and unmitigated mismatch and between the mitigated and partially mitigated mismatch. The partially mitigated plot is between the mitigated and unmitigated plots, but lies closer to the unmitigated plot which suggests that the effect of the mediation is not linear. The plot lines for both the mitigated and partially mitigated factors are very similar in shape with TR generally increasing slightly with disc size for the largest disc size.

6.7.2 Testing for significant effects

Significance testing is conducted to check for any effects of mismatch and disc size. Only the mitigated and unmitigated levels of the mismatch factor are tested. Mean TR values are subjected to a two-way repeated measures ANOVA.

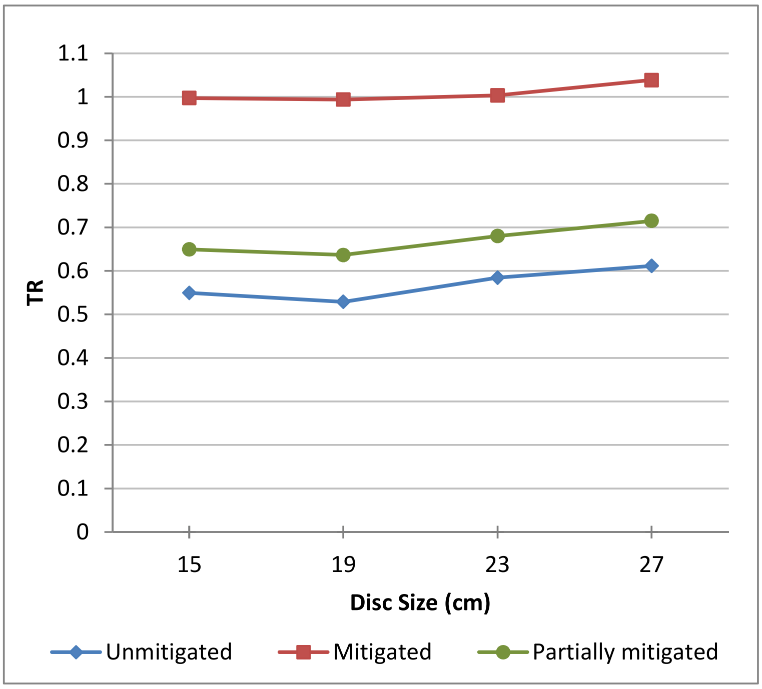
Tests for normality and outlier detection are conducted using standardised residuals of the data. TR is found to be normally distributed for all group combinations of mismatch and levels of disc size, as assessed by Shapiro-Wilk's test (p > .05). One outlier is detected, as assessed by box plot. The decision is made to keep the value even though it is extreme as for that group of results it is consistent with that individual's results across the other groups.

For the interaction between mismatch and size on PR, Mauchly's test indicates that the assumption of sphericity has been violated, $\chi^2(5) = 16.890$, $p = .005$, and therefore the Greenhouse-Geisser correction is applied ($\varepsilon = .528$). There is no significant interaction between mismatch and disc size, $F(1.584, 20.598) = .687$, $p > .05$.

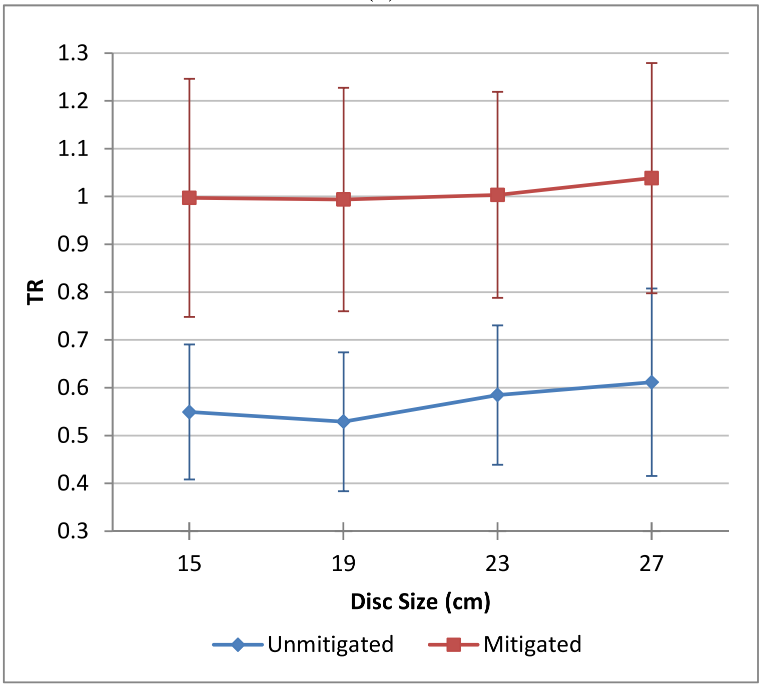
There is significant main effect of mismatch, $F(1, 13) = 41.126$, $p < .001$.

For the disc size factor, Mauchly's test indicates that the assumption of sphericity has been violated, χ^2 $(5) = 23.363$, $p = .0$, and therefore the Greenhouse-Geisser correction is applied ($\varepsilon = .464$). There is no significant main effect of disc size, $F(1.393, 18.113) = 3.460$, $p > .05$.

These statistical tests were repeated with the removal of the participant data that contained the outlier; the overall result did not change.



(a)



(b)

Fig. 6.6 (a) Plots of the estimated marginal means of the mitigated, unmitigated and partially mitigated mismatch conditions. Errors bars omitted for readability. (b) Plots of the estimated marginal means of the mitigated and unmitigated mismatch conditions. The error bars denote $\pm\sigma_t$, the standard deviations.

Table 6.3 Slope, intercept, coefficient of determination (R^2) and p-value of **mitigated** accommodation-vergence mismatch versus participant age, for each disc size.

Disc size	Slope	Intercept	R^2	p
15cm	0.0102	0.6227	0.1525	$> .10$
19cm	0.0100	0.6276	0.1658	$> .10$
23cm	0.0083	0.7013	0.1326	$> .10$
27cm	0.0051	0.8504	0.0412	$> .10$

Table 6.4 Slope, intercept, coefficient of determination (R^2) and p-value of **unmitigated** accommodation-vergence mismatch versus participant age, for each disc size.

Disc size	Slope	Intercept	R^2	p
15cm	-0.0038	0.6878	0.0315	$> .10$
19cm	-0.0034	0.6525	0.0208	$> .10$
23cm	-0.0002	0.5906	0.0000	$> .10$
27cm	-0.0016	0.6695	0.0028	$> .10$

6.7.3 Examining relationship between TR and age

The correlation between TR and age is investigated by fitting a straight line. Two plots — mitigated and unmitigated accommodation-vergence mismatch condition. The results are displayed in Tables 6.3 and 6.4, the corresponding plots in Figure 6.7.

Table 6.5 Slope, intercept, coefficient of determination (R^2) and p-value of **mitigated** accommodation-vergence mismatch versus participant age, for each disc size. This data is a subset of the results within the participant age range of **23–46**.

Disc size	Slope	Intercept	R^2	p
15cm	0.0264	0.1081	0.4368	$< .5$
19cm	0.0255	0.1326	0.4791	$< .5$
23cm	0.0222	0.2568	0.4162	$< .5$
27cm	0.0202	0.3713	0.2832	$< .10$

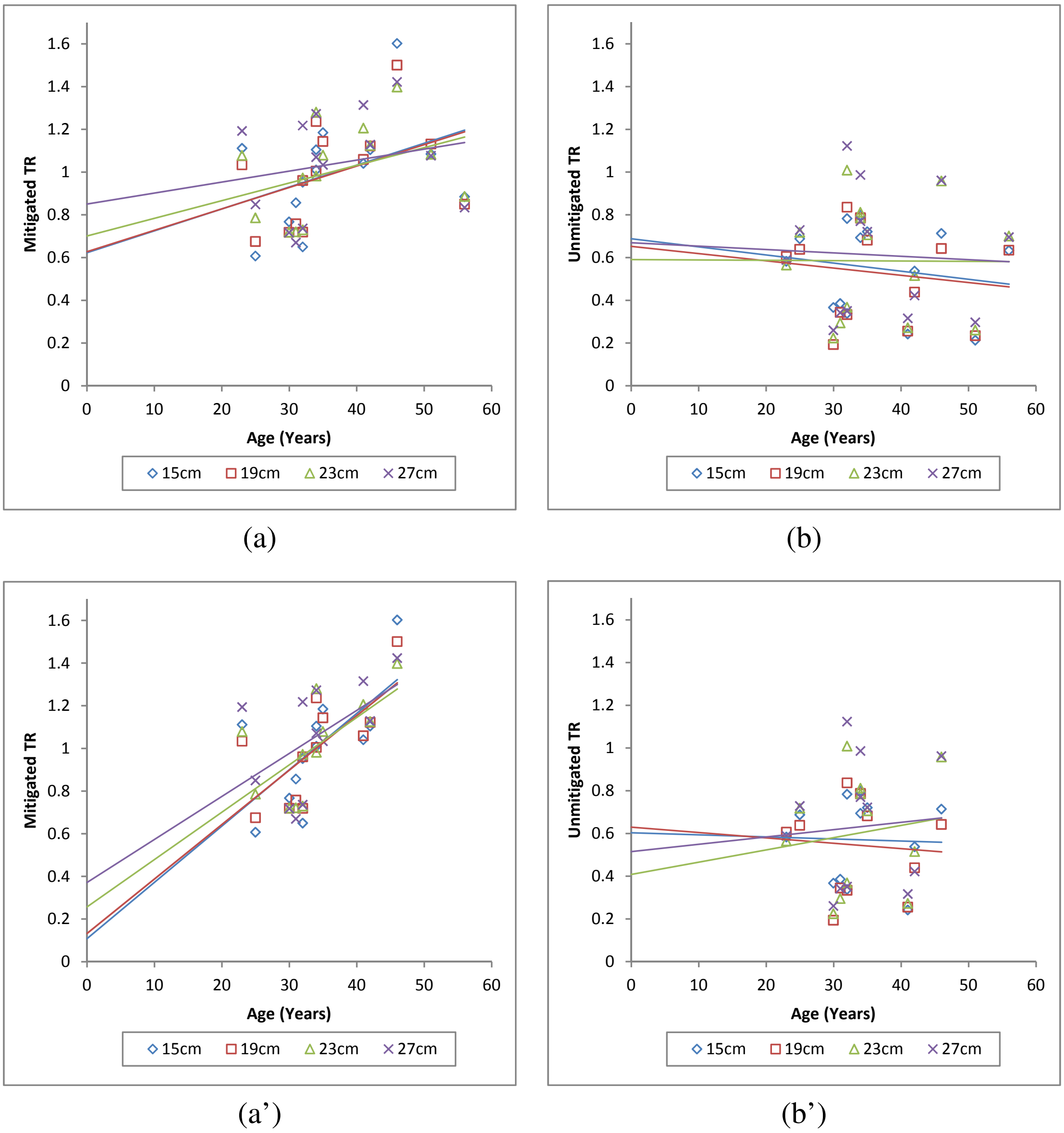


Fig. 6.7 Correlation of TR values and age for each disc size for both (a) Mitigated and (b) Unmitigated accommodation-vergence mismatch conditions. Plots (a') and (b') use a subset of the same data, but 2 participants with ages > 46 are omitted. Model-fitting parameters, weighted R^2 and p values can be found in Tables 6.3, 6.4, 6.5 and 6.6.

Table 6.6 Slope, intercept, coefficient of determination (R^2) and p-value of **unmitigated** accommodation-vergence mismatch versus participant age, for each disc size. This data is a subset of the results within the participant age range of **23–46**.

Disc size	Slope	Intercept	R^2	p
15cm	-0.0010	0.6029	0.0011	$> .10$
19cm	-0.0025	0.6294	0.0057	$> .10$
23cm	-0.0057	0.4082	0.0202	$> .10$
27cm	-0.0034	0.5148	0.0060	$> .10$

6.8 Discussion

As discussed in the previous section, the ANOVA found that there is a statistically significant difference between the mitigated and unmitigated mismatch conditions. The direction of movement of TR values is similar to the findings of Thouless (1932), in the sense that the reduction in the quality of stereo cues resulted in a drop of TR across participants. However, the regression analysis reveals no correlation between the mismatch conditions, so the relationship between them is not simple to determine.

It is further found that the disc size is no longer a factor, which may be because the variable ambient light levels were adequately controlled by using grey rather than white discs.

The range of TR values is similar in both experiments, but it was expected that mean TR values in the mitigated condition would resemble those of the physical condition in the previous experiment. However, the TR values are close to unity in this experiment. Although it is possible that they might have unknowingly reverted to doing objective matching, this seems unlikely given the care taken in instructing them.

Another possibility is that the change in focus of the screen had an effect on the participants. At the mitigated position the focal distance to the standard disc is 2m and it is 1m to the response disc. Similarly, for the unmitigated position, the focal distance to the standard disc is 2m and it is 2.72m to the response disc. The change in focal distance when looking back and forth between the response and standard discs at the closer focal distances of the mitigated position is potentially more strenuous on the ciliary muscles of the eyes. This theory may be supported by comments made by participants after the trials. The group of participants that performed the unmitigated trial block first, commented that the perspective matching task suddenly 'felt' harder to do in the mitigated trial block, whereas participants switching from the mitigated to unmitigated trial blocks suddenly 'felt' as though the task was easier.

It is also possible that, though measures were taken to reduce the effect of ghosting, the result was due to the presence of crosstalk in the unmitigated condition.

As with the previous experiment the relationship between TR and participant age was examined to observe any hint of ageing affecting phenomenal regression (Figure 6.7). The results in Tables 6.3 and 6.4 again reveal no such correlation. However, taking into account the effect of Presbyopia, the results were re-analysed, but 2 participants with ages > 46 were omitted leaving a subset of ages 23–46; see plots (Figures 6.7 (a') and (b')). The results for the mitigated disc (Table 6.5) indicate that PR increases with participant age. However, the unmitigated results (Table 6.6) are uncorrelated. This suggests that the unmitigated condition was not veridical to a real-world viewing condition.

6.9 Key property of TR established

The goal of this experiment was to prove the research hypothesis by establishing TR to be responsive, in addition to the established key properties of validity and reliability.

The result indicates that TR is responsive to changes in stereo cue conditions in a VE. This chapter thus provides the additional fulfilment to objective [iii] of the research aim by investigating PR as a VE assessment metric.

Chapter 7

Conclusion

7.1 Summary of the thesis

The general aim of the research, as first discussed in Chapter 1, was to identify a direct measure of VE quality that relies on participants' innate perceptual response to visual stimuli. The three objectives, that were identified in Chapter 1 to fulfil the research aim are as below:

i. Identify a key element of spatial perception where there is a suspected deficit in the VE compared to the real world.

Chapter 2 fulfilled research objective [i] by reviewing the body of work relating to the deficit of depth perception in a VE. The chapter also defines the scope of the experimental work as occurring within MR space based on the classification of the different kinds of VE spaces, and identifies the technological scope of the research as large-screen immersive VEs.

ii. Identify a directly measurable property capable of characterising subjective perceptual experience of visual stimuli.

Chapter 3 fulfilled research objective [ii] by investigating the psychological trait PR and its numerical measure TR. The properties that make PR interesting for use in VEs are as below:

- It is influenced by perceptual indicators like visual depth cues
- It responsive to changes in stereo cue conditions
- It is an innate, involuntary and unmodifiable response to visual stimuli
- It is not mediated by experience or skill or previous training, i.e. participants cannot get better by practice

- It is subjective, showing individual differences due to traits of a person's psychology.

Having fulfilled objectives [i] and [ii], objective [iii] was refined and restated in Chapter 4.

iii. Within the scope of depth perception for large-screen immersive display VE, investigate the performance of TR to establish its value as a VE assessment metric.

The hypothesis of the research was stated as: *PR is a potentially useful measure of veridical perception to compare visual responses to virtual and physical stimuli*. In order to fulfil objective [iii] and to prove the hypothesis, key properties of TR were identified to be established. The three properties that are first identified in Chapter 4 are that TR needs to proven to be valid (measure the perceptual deficit), reliable (produce consistent results when there is no evidence of change), and responsive (detect change).

By conducting an experiment comparing TR between virtual and physical stimuli, Chapter 5 established that TR is a valid measure, having observed that the effect, PR, occurs in both the real world and a VE. The results also indicate that TR is reliable for the duration of the experiment, the details of which can be found in sections 5.7 and 5.8. It also verified that TR varies between individuals.

The second experiment, described in Chapter 6, investigated the effect of mitigating accommodation-vergence mismatch of stimuli in order to test for the responsiveness of TR to changes in depth cue conditions in a VE. TR was shown to be responsive to changes in stereo conditions in a VE by the statistically significant result discussed in section 6.7.

Having fulfilled all three objectives of the research and demonstrating that the three key properties of validity, reliability and responsiveness are fulfilled by TR, the hypothesis of the research can be considered to be proven. PR is, thus, a potential metric of veridical perception in VEs.

7.2 Contributions to the field

7.2.1 PR

The work has clarified the terminology, verified that the effect occurs and verified that PR, as reported by Thouless, is indeed subjective with its affect diminished by reduction in stereo conditions.

This research has rendered the subject of PR accessible for those outside the psychology community. The informed re-telling of Thouless' work clarifies the rationale behind the

TR equation and the methodology for measuring perspective size. The lesson from the exploratory pilots and subsequent follow up experiments provide guidelines to designing and conducting such studies. The research highlights how important it is to devise clear participant instruction for perspective matching.

We further related the effect of PR to VE depth perception showing that PR is not just a study of size constancy.

PR is the modification of the perception of retinal stimulus, to the extent that subjects cannot accurately estimate perspective size. It is not emphasised in the literature just how remarkable this effect of PR is. It can be speculated that this is due to an incomplete understanding of the effect, given that we found the need to perform the task in order to experience it.

7.2.2 PR in VE

This is the first study of PR in VE and the results indicate that the effect of PR occurs in VE as it does in the real world.

The results have verified that TR is a valid measure of PR of perspective size for both physical and virtual stimuli and that it is responsive to changes in focal cue conditions in a VE.

An individual's TR has been demonstrated to have temporal stability for the duration of the testing period and, therefore, is reliable measure.

This research has drawn attention to a method to measure subjective or individual experience of visual stimuli in a VE that does not rely on the skill of the participant conducting the task. In contrast to other groups' work, PR is not an action based measure that has been guided by judgement. It is instead a direct measurement of characteristics of visual stimuli. PR is an innate, involuntary and unmodifiable response to visual stimuli. PR is not mediated by experience, skill or previous training. It cannot be consciously manipulated as long as the participant performs the matching task given to them, as instructed and with both eyes fully open and focussing on each stimuli.

7.3 Future work

7.3.1 Key properties

Two of the key properties require further investigation. The first property that requires further investigation is the responsiveness of TR. Although it has been established that TR is responsive to a large change in VE stereo conditions, how sensitive it is remains to be established. If TR is not responsive to small to medium changes in VE conditions its usefulness would be limited. The second property of TR that needs further investigation is the reliability of PR. Although PR was shown to be stable for the duration of the experiments, it is not known how robust it is over a long period of time. Thouless anecdotally reported that he thought it was robust for a period of 24 hours as some participants were tested on subsequent days and exhibited similar TR values. However, this was not satisfactorily verified.

7.3.2 Relation to other studies

The research has established that PR and the use of TR to obtain a metric of VE quality has potential. In Chapter 1, training was highlighted as a VE application domain that would benefit from a VE quality metric. TR has the potential to be used as a parameter of any mathematical function of transfer of training. However, further work is still required.

The reason we chose to study PR of size was because it has the linked quality of perception of distance. It would be useful to investigate if individual TR correlates with errors in distance perception.

7.3.3 Suggestion of other experiment designs

The focus of this research was perspective size, since this involves the linked character of distance, which has been much-studied in VEs. PR of the other characteristics investigated by Thouless like PR of brightness, colour and perspective shape would be interesting to investigate in VEs. For instance, testing PR of perspective shape could prove useful to investigate techniques used to implement stereo viewing frusta in a VE.

One question remaining from the experiment detailed in Chapter 5 is whether the presence of the real physical tripod beneath the virtual stimulus did mitigate the accommodation-vergence mismatch. This can be investigated further now, since the recent addition of floor projection in HIVE makes it possible to replace the physical tripod with a virtual tripod.

One of the questions that remains to be answered based on the experiment described in Chapter 6, is that of how sensitive TR is to smaller changes in stereo conditions. The separation of the mitigated and unmitigated conditions was the largest feasible with the equipment available. In order to establish the sensitivity of TR to stereo conditions we can propose an experiment to slowly mitigate the accommodation-vergence mismatch using a psychophysical testing methodology.

References

Adams, R. M., Stancampiano, B., McKenna, M., and Small, D. (2002). Case study: A virtual environment for genomic data visualization. In *Proceedings of the Conference on Visualization '02*, VIS '02, pages 513–516, Washington, DC, USA. IEEE Computer Society.

Adelstein, B. D., Lee, T. G., and Ellis, S. R. (2003). Head Tracking Latency in Virtual Environments: Psychophysics and a Model. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, pages 2083–2087.

Aggarwal, R., Ward, J., Balasundaram., I., Sains, P., Athanasiou, T., and Darzi, A. (2007). Proving the effectiveness of virtual reality simulation for training in laparoscopic surgery. *Ann Surg*, 246(5):771–9.

Allison, R. S., Harris, L., Jenkins, M., Jasopbedzka, U., and J.E., Z. (2001). Tolerance of temporal delay in virtual environments. In *IEEE Virtual Reality 2001*, pages 247–254. IEEE Comput. Soc.

Azevedo, A, J. J. . C. P. (2014). Combining eeg data with place and plausibility responses as an approach to measuring presence in outdoor virtual environments. *Presence: Teleoperators & Virtual Environments*, 23(4):354–368.

Baldwin, T. T. and Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 41(1):63–105.

Benford, S., Brown, C., Reynard, G., and Greenhalgh, C. (1996). Shared spaces: Transportation, artificiality, and spatiality. In *CSCW'96*.

Benford, S., Greenhalgh, C., Reynard, G., Brown, C., and Koleva, B. (1998). Understanding and constructing shared spaces with mixed-reality boundaries. *ACM Transactions on Computer-Human Interaction*, 5:3(3):185–223.

Bossard, C., Kermarrec, G., Buche, C., and Tisseau, J. (2008). Transfer of learning in virtual environments: a new challenge? *Virtual Reality*, 12(3):151–161.

Braisby, N. and Gellatly, A. (2005). *Cognitive Psychology*. Oxford:Oxford University Press.

BSM (2008). British school of motoring (bsm) products, the simuator. Website at: www.autotrader.co.uk/CARS/buying/simulator.jsp. Last Accessed 17/08/10.

Charman, W., N. (2008). The eye in focus: accomodation and presbyopia. *Clin Exp Optom*, 91:3:207–225.

Chastine, J., Zhu, Y., Brooks, J., Owen, G., Harrison, R., and Weber, I. (2005). A collaborative multi-view virtual environment for molecular visualization and modeling. In *Coordinated and Multiple Views in Exploratory Visualization, 2005. (CMV 2005). Proceedings. Third International Conference on*, pages 77–84.

Coren, S. (2003). *Handbook Of Psychology: History Of Psychology*, volume 1, chapter Sensation and Perception, pages 85–108. John Wiley & Sons.

Coren, S., Ward, L. M., and Enns, J. T. (2004). *Sensation and Perception*. Wiley, 6 edition.

Cottingham, J., Stoothoff, R., and Murdoch, D. (1988). *Selected Philosophical Writings of Descartes*, volume 1. Cambridge University Press. Date of original manuscript 1637.

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A., and Thompson, W. B. (2005). The influence of restricted viewing conditions on egocentric distance perception: Implications for real and virtual environments. *Perception*, 34:191–204.

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

Cutting, J. E. and Vishton, P. M. (1995). *Handbook of perception and cognition; Perception of space and motion*, volume 5, chapter Perceiving layout and knowing distances: The integration , relative potency, and contextual use of different information about depth, pages 69–117. San Diego, CA: Academic Press.

Ellis, S. (1991). Nature and origin of virtual environments: A bibliographic essay. *Computing Systems in Engineering*, 2(4):321–347.

Ellis, S. R. (1994). What are virtual environments? *IEEE Comput. Graph. Appl.*, 14(1):17–22.

Ellis, S. R. and Menges, B. M. (1998). Localization of virtual objects in the near visual field. *Hum. Factors*, 40:415–431.

Elner, K. (2007). The hive as an instrument of perceptual psychology. Msc, University of Hull.

Eysenck, M. and Keane, T. (2005). *Cognitive Psychology. A student handbook*. East Sussex: Psychology Press Ltd., 4 edition.

Field, P. (2012). Moving from military equipment to simulator enables significant cost savings. Web Article at: www.trainingandsimulationforum.net/news/industrynews/news/229#. Last Accessed 24/03/2014.

Friedman, D., Leeb, R., Guger, C., Steed, A., Pfurtscheller, G., and Slater, M. (2007). Navigating virtual reality by thought: What is it like? *Presence: Teleoper. Virtual Environ.*, 16(1):100–110.

Gamesutra (2013). Global games market grows 6 Website at: www.gamasutra.com/view/pressreleases/193010/Global_Games_Market_Grows_6_to_704bn_in_2013.php. Last Accessed 24/03/2014.

Garrett, A., Aguilar, M., and Barniv, Y. (2002). A recurrent neural network approach to virtual environment latency reduction. In *Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, volume 3, pages 2288–2292.

Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.

Gilinsky, A. S. (1955). The effect of attitude upon the perception of size. *The American Journal of Psychology*, 68(2):pp. 173–192.

Gottheil, E. and Bitterman, M. E. (1951). The measurement of shape-constancy. *The American Journal of Psychology*, 64:3:pp. 406–408.

Gregory, R. (1998). *Eye and Brain. The Psychology of Seeing*. Oxford University Press.

Haber, R. N. and Levin, C. A. (1989). *The moon illusion*, chapter The lunacy of moon watching: Some preconditions on explanations of the moon illusion, pages 299–318. Hillsdale, NJ: Erlbaum.

Haber, R. N. and Levin, C. A. (2001). The independence of size perception and distance perception. *Perception & Psychophysics*, 63(7):1140–1152.

Harding, C. (2004). Modeling geoscience data in a multisensory virtual environment. *Computing in Science Engineering*, 6(1):89–92.

Hershenson, M. (1999). *Visual Space Perception — A Primer*. The MIT Press.

Hoffman, H. G., Chambers, G. T., Meyer, Walter J., I., Arceneaux, L. L., Russell, W. J., Seibel, E. J., Richards, T. L., Sharar, S. R., and Patterson, D. R. (2011). Virtual reality as an adjunctive non-pharmacologic analgesic for acute burn pain during medical procedures. *Annals of Behavioral Medicine*, 41(2):183–191.

Hofstetter, H., W. (1950). A useful age-amplitude formula. *Optom World*, 38:42–45.

Holway, A. and Boring, E. (1941). Determinants of apparent visual size with distance variant. *American Journal of Psychology*, 54:121–151.

IKEA (2013). Place ikea furniture in your home with augmented reality. Website at: www.youtube.com/watch?v=vDNzTasuYEwt=72. Last Accessed 24/03/2014.

IKEA (2014). Ikea catalogue 2014. Website at: www.ikea.com/ms/en_GB/virtual_catalogue/online_catalogues.html. Last Accessed 24/03/2014.

Interrante, V., Ries, B., and Anderson, L. (2006). Distance perception in immersive virtual environments, revisited. In *Proceedings of the IEEE conference on Virtual Reality*, VR '06, pages 3–10, Washington, DC, USA. IEEE Computer Society.

Interrante, V., Ries, B., Lindquist, J., Kaeding, M., and Anderson, L. (2008). Elucidating factors that can facilitate veridical spatial perception in immersive virtual environments. *Presence: Teleoper. Virtual Environ.*, 17(2):176–198.

Jerald, J., Peck, T., Steinicke, F., and Whitton, M. (2008). Sensitivity to scene motion for phases of head yaws. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization*, APGV '08, pages 155–162, New York, NY, USA. ACM.

Jerald, J. and Whitton, M. (2009). Relating scene-motion thresholds to latency thresholds for head-mounted displays. In *Virtual Reality Conference, 2009. VR 2009. IEEE*, pages 211–218.

Johansson, G. (1975). *Visual Motion Perception*. USA: Scientific America, Inc.

Joynson, R. B. (1949). The problem of size and distance. *Quarterly Journal of Experimental Physiology*, 1:3:pp. 119–135.

Juan, M., Alcaniz, M., Monserrat, C., Botella, C., Banos, R., and Guerrero, B. (2005). Using augmented reality to treat phobias. *Computer Graphics and Applications, IEEE*, 25(6):31–37.

Juan, M. C. and Pérez, D. (2009). Comparison of the levels of presence and anxiety in an acrophobic environment viewed via hmd or cave. *Presence: Teleoper. Virtual Environ.*, 18(3):232–248.

Judd, C. H. (1908). The relation of special training to general intelligence. *Educ. Rev.*, 36:28–42.

Jung, J. Y., Adelstein, B. D., and Ellis, S. R. (2000). Discriminability of prediction artifacts in a time-delayed virtual environment. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 445, pages 499–502.

Kenyon, R. V., Sandin, D., Smith, R. C., Pawlicki, R., and Defanti, T. (2007). Size-constancy in the cave. *Presence: Teleoper. Virtual Environ.*, 16(2):172–187.

Kickstarter (2014). Kickstarter. Website available at www.kickstarter.com/. Last accessed 31/03/14.

Klein, E., Swan, J. E., Schmidt, G. S., Livingston, M. A., and Staadt, O. G. (2009). Measurement protocols for medium-field distance perception in large-screen immersive displays. In *Proceedings of the 2009 IEEE Virtual Reality Conference*, VR '09, pages 107–113, Washington, DC, USA. IEEE Computer Society.

Klimenko, S., Nielson, G. M., Nikitina, L., and Nikitin, I. (2002). Adventures of mobius band: Mathematical visualization in virtual environment. In *Proceedings of the Eighteenth Annual Symposium on Computational Geometry*, SCG '02, pages 281–282, New York, NY, USA. ACM.

Knapp, J. M. and Loomis, J. M. (2004). Limited field of view of head-mounted displays is not the cause of distance underestimation in virtual environments. *Presence: Teleoper. Virtual Environ.*, 13(5):572–577.

Kuhl, S. A., Thompson, W. B., and Creem-Regehr, S. H. (2009). Hmd calibration and its effects on distance judgments. *ACM Trans. Appl. Percept.*, 6(3):19:1–19:20.

Kunz, B. R., Wouters, L., Smith, D., Thompson, W. B., and Creem-Regehr, S. H. (2009). Revisiting the effect of quality of graphics on distance judgments in virtual environments: A comparison of verbal reports and blind walking. *Attention, Perception, & Psychophysics*, 71:1284–1293.

Lear, A. (1997). Virtual reality provides real therapy. *Computer Graphics and Applications, IEEE*, 17(4):16–20.

Lee, D. (2014). Facebook buys virtual reality headset start-up for $2bn. Website at: www.bbc.co.uk/news. Last Accessed 31/03/2014.

Lo, C. Y., Chao, Y. P., Chou, K. H., Guo, W. Y., Su, J. L., and Lin, C. P. (2007). Dti-based virtual reality system for neurosurgery. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 1326–1329.

Loomis, J. M. and Knapp, J. M. (2003). *Virtual and Adaptive Environments Applications, Implications, and Human Performance Issues*, chapter Visual Perception of Egocentric distance in Real and Virtual Environments, pages 21–46. CRC Press.

Maersk (2014). Maersk: Drilling simulator. Website at: www.maerskdrilling.com/AboutUs/DrillingSimulator/Pages/drilling-simulator.aspx. Last Accessed 24/03/2014.

Mainwaring, J. (2012). Eyesim: The future of training is virtual reality. Web Article at: www.rigzone.com/news/oil_gas/a/122264/EYESIM_The_Future_of_Training_is_Virtual_Reality. Last Accessed 24/03/2014.

Marr, D. (1982). *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman.

Mather, G. (2009). *Foundations of Sensation and Perception*. Psychology Press, 2 edition.

McCullum, K. (2011). The benefits of helicopter door gunnery simulation. *MT2*, 16(5).

McDonald, R. P. (1962). An artifact of the brunswick ratio. *The American Journal of Psychology*, 75:1:pp. 152–154.

Meehan, M., Razzaque, S., Whitton, M. C., and Brooks, Jr., F. P. (2003). Effect of latency on presence in stressful virtual environments. In *Proceedings of the IEEE Virtual Reality 2003*, VR '03, pages 141–, Washington, DC, USA. IEEE Computer Society.

Milgram, P. and Kishino, F. (1994). A taxonomy of mixed reality visual displays. *IEICE TRANS. INF. & SYST.*, E77-D(12).

Mohler, B. J., Bülthoff, H. H., Thompson, W. B., and Creem-Regehr, S. H. (2008). A full-body avatar improves egocentric distance judgments in an immersive virtual environment. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization*, APGV '08, pages 194–, New York, NY, USA. ACM.

Mohler, B. J., Creem-Regehr, S. H., Thompson, W. B., and Bülthoff, H. H. (2010). The effect of viewing a self-avatar on distance judgments in an hmd-based virtual environment. *Presence: Teleoper. Virtual Environ.*, 19(3):230–242.

Muller, P., Cohn, L. J., Schmorrow, C. D., Stripling, R., Stanney, K., Milham, L., Jones, D., Whitton, M. C., and Fowlkes, J. E. (2006). The fidelity matrix: Mapping system fidelity to training outcome. In *The Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*, volume 2006. NTSA.

Murgia, A. and Sharkey, P. M. (2009). Estimation of distances in virtual environments using size constancy. *The International Journal of Virtual Reality*, 8(1):67–74.

Napieralski, P. E., Altenhoff, B. M., Bertrand, J. W., Long, L. O., Babu, S. V., Pagano, C. C., Kern, J., and Davis, T. A. (2011). Near-field distance perception in real and virtual environments using both verbal and action responses. *ACM Trans. Appl. Percept.*, 8(3):18:1–18:19.

OculusVR (2012). Oculus rift: Step into the game. Kickstarter crowdfunding platform at www.kickstarter.com. Last Accessed 31/03/2014.

OculusVR (2014). Oculus vr. Website at: www.oculusvr.com/. Last Accessed 24/03/2014.

O'Toole, R., V., Playter, R., R., Krummel, T., M., Blank, W., C., Cornelius, N., H., Roberts, W., R., Bell, W., J., and Raibert, M. (1999). Measuring and developing suturing technique with a virtual reality surgical simulator. *J Am Coll Surg*, 189(1):114–27.

Palmer, S., E. (2002). *Vision Science — Photons to Phenomenology*. The MIT Press.

Papadakis, G., Mania, K., Coxon, M., and Koutroulis, E. (2011). The effect of tracking delay on awareness states in immersive virtual environments: An initial exploration. In *Proceedings of the 10th International Conference on Virtual Reality Continuum and Its Applications in Industry*, VRCAI '11, pages 475–482, New York, NY, USA. ACM.

Pennington, N., Nicolich, R., and Rahm, J. (1995). Transfer of training between cognitive subskills: Is knowledge use specific? *Cognitive Psychology*, 28(2):175 – 224.

Philipp, M. C., Storrs, K. R., and Vanman, E. J. (2012). Sociality of facial expressions in immersive virtual environments: A facial {EMG} study. *Biological Psychology*, 91(1):17 – 21.

Piryankova, I. V., de la Rosa, S., Kloos, U., Bulthoff, H. H., and Mohler, B. J. (2013). Egocentric distance perception in large screen immersive displays. *Displays*, (0).

Pivkin, I., Yang, N., Richardson, P., Karniadakis, G., and Laidlaw, D. (2004). Visualization of blood platelets in a virtual environment. In *ACM SIGGRAPH 2004 Posters*, SIGGRAPH '04, pages 111–, New York, NY, USA. ACM.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., editors (1992). *Numerical Recipes in C: The Art of Scientific Computing*, chapter Modelling of Data, pages 666–668. Cambridge University Press, 2 edition.

QinetiQPlc (2013). Helicopter simulation training. Website at: www.qinetiq.com/what/capabilities/training-simulation/Pages/helicopter-simulators.aspx. Last Accessed 24/03/2014.

Ries, B., Interrante, V., Kaeding, M., and Anderson, L. (2008). The effect of self-embodiment on distance perception in immersive virtual environments. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology*, VRST '08, pages 167–170, New York, NY, USA. ACM.

Rizzo, A., Buckwalter, J. G., John, B. S., Newman, B., Parsons, T. D., Kenny, P. G., and Williams, J. (2012). STRIVE: stress resilience in virtual environments: A pre-deployment VR system for training emotional coping skills and assessing chronic and acute stress responses. *Medicine Meets Virtual Reality*, 19:379–385.

Rolland, J. P. and Fuchs, H. (2000). Optical versus video see-through head-mounted displays in medical visualization. *Presence: Teleoper. Virtual Environ.*, 9(3):287–309.

Rose, F., Attree, E., Brooks, B., Parslow, D., Penn, P., and Ambihaipahan, H. (2000). Training in virtual environments: transfer to real world tasks and equivalence to real task training. *Ergonomics*, 43(4):494–511.

Ruddle, R. and Lessels, S. (2006). Three levels of metric for evaluating wayfinding. *Presence: Teleoperators & Virtual Environments*, 15(6):637–654.

Sahm, C. S., Creem-Regehr, S. H., Thompson, W. B., and Willemsen, P. (2005). Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Trans. Appl. Percept.*, 2(1):35–45.

Sedgwick, H. A. (1986). *Handbook of Perception and Human Performance*, volume 1, chapter Space perception, pages 129–158. New York: John Wiley and Sons.

Singh, G., Swan, II, J. E., Jones, J. A., and Ellis, S. R. (2010). Depth judgment measures and occluding surfaces in near-field augmented reality. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, APGV '10, pages 149–156, New York, NY, USA. ACM.

Slater, M., Guger, C., Edlinger, G., Leeb, R., Pfurtscheller, G., Antley, A., Garau, M., Brogni, A., Steed, A., and Friedman., D. (2006). Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence : Teleoper. Virtual Environments*, 15(5):553–569.

Slater, M., Pertaub, D., and Steed, A. (1999). Public speaking in virtual reality: facing an audience of avatars. *Computer Graphics and Applications, IEEE*, 19(2):6–9.

Sohn, H., Jung, Y. J., and Ro, Y. M. (2014). Crosstalk reduction in stereoscopic 3d displays: Disparity adjustment using crosstalk visibility index for crosstalk cancellation. *Opt. Express*, 22(3):3375–3392.

Stanney, K. (1995). Realizing the full potential of virtual reality: human factors issues that could stand in the way. In *VRAIS '95: Proceedings of the Virtual Reality Annual International Symposium (VRAIS'95)*, pages 28+, Washington, DC, USA. IEEE Computer Society.

Strickland, D. and Chartier, D. (1997). Eeg measurements in a virtual reality headset. *Presence: Teleoperators & Virtual Environments*, 6(5):581.

TheGadgetShow (2011). Up close with the battlefield 3 simulator! Website at: gadgetshow.channel5.com/gadget-show/gadget-news/up-close-with-the-battlefield-3-simulator. Last Accessed 24/03/2014.

Thomas, B. H. (2012). A survey of visual, mixed, and augmented reality gaming. *ACM Comput. Entertain*, 10(3).

Thompson, W. B., Willemsen, P., Gooch, A. A., Creem-Regehr, S. H., Loomis, J. M., and Beall, A. C. (2004). Does the quality of the computer graphics matter when judging distances in visually immersive environments. *Presence: Teleoper. Virtual Environ.*, 13(5):560–571.

Thorndike, E. L. and Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other influence. *Psychological Review*, 8:247–261.

Thouless, R. H. (1931a). Phenomenal regression to the real object. i. *British Journal of Psychology. General Section*, 21(4):339–359.

Thouless, R. H. (1931b). Phenomenal regression to the real object. ii. *British Journal of Psychology. General Section*, 22(1):1–30.

Thouless, R. H. (1932). Individual differences in phenomenal regression. *British Journal of Psychology. General Section*, 22(3):216–241.

Thouless, R. H. (1933). A racial difference in perception. *The Journal of Social Psychology*, 4(3):330–339.

Tsirlin, I., W. L. M. and Allison, R. (2011). The effect of crosstalk on the perceived depth from disparity and monocular occlusions. *Broadcasting, IEEE Transactions on*, 57(2):445–453.

Tsirlin, I., W. L. M. and Allison, R. (2012). Crosstalk reduces the amount of depth seen in 3d images of natural scenes. In *Proceedings of the Stereoscopic Displays and Applications XXII Conference, San Francisco*.

Tzafestas, C., Birbas, K., Koumpouros, Y., and Christopoulos, D. (2008). Pilot evaluation study of a virtual paracentesis simulator for skill training and assessment: The beneficial effect of haptic display. *Presence: Teleoperators & Virtual Environments*, 17(2):212–229.

Van Baren, J. and IJsselsteijn, W. (2004). Measuring presence: A guide to current measurement approaches. Technical Report 5. Deliverable of the OmniPres project IST-2001-39237.

Vertual (2014). Vertual - innovative tools for training and education. Website at: www.vertual.eu/home/. Last Accessed 24/03/2014.

VirtualAerospace (2014). Virtual aerospace flight simulators. Website at: www.virtual-aerospace.com/. Last Accessed 24/03/2014.

VirtualAviationLtd. (2013). Flight simulator training. Website at: www.virtualaviation.co.uk/pages/flight-simulator-training. Last Accessed 01/01/2013.

Willemsen, P., Colton, M. B., Creem-Regehr, S. H., and Thompson, W. B. (2009). The effects of head-mounted display mechanical properties and field of view on distance judgments in virtual environments. *ACM Trans. Appl. Percept.*, 6(2):8:1–8:14.

Willemsen, P. and Gooch, A. A. (2002). An experimental comparison of perceived egocentric distance in real, image-based, and traditional virtual environment using direct walking tasks. Technical report, University of Utah Computer Science.

Willemsen, P., Gooch, A. A., Thompson, W. B., and Creem-Regehr, S. H. (2008). Effects of stereo viewing conditions on distance perception in virtual environments. *Presence: Teleoper. Virtual Environ.*, 17(1):91–101.

Witmer, B. and Singer, M. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240.

Woods, A. J. (2011). How are crosstalk and ghosting defined in the stereoscopic literature?

Yuan, M., Khan, I., Farbiz, F., Yao, S., Niswar, A., and Foo, M. H. (2013). A mixed reality virtual clothes try-on system. *Multimedia, IEEE Transactions on*, 15(8):1958–1968.

Yue, L., Yongtian, W., Yu, L., Jinchao, L., and Liang, L. (2006). Key issues for ar-based digital reconstruction of yuanmingyuan garden. *Presence: Teleoperators & Virtual Environments*, 15(3):pp. 336–340. Academic Search Premier, EBSCOhost, viewed 14 April 2014.

Zhang, S., Demiralp, C., Zhang, S., Demiralp, ., Keefe, D. F., Dasilva, M., Basser, P., Pierpaoli, C., Chiocca, E., Deisboeck, T., Laidlaw, D. H., and Greenberg, B. D. (2001). An immersive virtual environment for dt-mri volume visualization applications: a case study. In *In Proceedings of IEEE Visualization 2001*, pages 437–440.

Ziegeler, S., Moorhead, R. J., Croft, P. J., and Lu, D. (2001). The metvr case study: Meteorological visualization in an immersive virtual environment. In *Proceedings of the Conference on Visualization '01*, VIS '01, pages 489–492, Washington, DC, USA. IEEE Computer Society.