Fidelity Metrics for Virtual Environment Simulations based on Human Judgements of Spatial Memory Awareness States

Aikaterini Mania

A thesis submitted to the University of Bristol, UK in accordance with the requirements for the degree of Doctor of Philosophy in the Faculty of Engineering, Department of Computer Science.

2001

Abstract

The aim of interactive computer graphics scenes and Virtual Environment (VE) technologies for simulation and training is the creation of accurate, high-quality imagery and interaction interfaces that faithfully represent a real-world task situation. Reliable fidelity evaluation techniques are essential in order to assess relevant implementations. A commonly employed strategy is to compare task performance in the VE in relation to the real world scene being represented. Spatial perception tasks are often incorporated in benchmarking processes as such, since spatial awareness is crucial for human performance efficiency. Such approaches, however, are not sufficient to assess the fidelity of VE systems. They are often limited to a specific application and are not based on formal frameworks but in most cases on arbitrary selected spatial perception tasks. This thesis introduces a metric, based on human judgements of spatial memory awareness states for assessing the simulation fidelity of a VE in relation to its real scene counterpart. This framework is based on the cognitive processes participants' employ in order to retrieve the memory of a space. Participants could describe how they make their spatial recollections by selecting between four choices of awareness states. These depend on the level of visual mental imagery involved during retrieval, the familiarity of the recollection and also include guesses, even if informed. In order to demonstrate the differences between using task performance based metrics and human evaluation of cognitive awareness states, a set of VEs displayed mainly on Head Mounted Displays (HMDs) were created. Resulting scenes were then compared to the real task situation they represented by employing the spatial memory awareness states methodology as well as assessments of presence, simulator sickness and responses to lighting. The experimental results are presented in this thesis, with an emphasis on probability-based formal analysis, revealing a variation of the distribution of participants' awareness states across conditions, especially when task performance failed to reveal any. Simulation of task performance does not necessarily lead to simulation of the cognitive processes employed in order to complete the task at hand for VE display technologies. The general premise of this thesis is focusing on *'how'* tasks are achieved, rather than only as in earlier VE simulation research, on *'what'* is achieved.

Declaration

The work in this thesis is original and no portion of the work referred to here has been submitted in support of an application for another degree or qualification of this or any other university or institution of learning.

Signed: Date:

Aikaterini Mania

Acknowledgements

I would like to thank Dr. Alan Chalmers for his continuous encouragement and enthusiasm, Prof. Tom Troscianko for his invaluable comments and also Prof. Martin Conway for his help with memory. Thanks are also due to all my colleagues and friends in the Department of Computer Science in Bristol who have helped me along the way, Angus, Mark (thanks for your help), Erik (thanks for proof-reading!), Jimi, Henk, Cliff (thanks for the digital compass), John, Oli, Sarah, Ki, Kate, the whole 1.15 and vision lab, dearest Claire, Ann Prowse and everybody that has patiently participated in pilot studies and experiments, I am grateful. Thank you to Greg and Holly. Thank you also to all my colleagues in the University of Sussex for giving me the time and peace of mind to actually finish my thesis, I appreciate their patience.

A special thank you to Hewlett Packard Laboratories and everybody involved there for their generous funding that allowed me to travel a lot in conferences and also for the continuous access to their facilities and my cubicle for the last five years. Thank you to Dr. Catherine Slater, my boss when I was a full-time researcher at Hewlett Packard Laboratories who gave me the freedom to explore any research ideas and also thank you to Dr. Jeremy Gunawardena, my initial contact with HP for his help. A ‘big’ thank you goes to Dr. Rycharde Hawkes for all his support.

Thanks to Ann for all the support, the jokes and the brilliant times in conferences (we got it covered, I remember!). Thank you to Ella for all her support as well. Thank you to my precious friend Katia from Greece, to Manolis and Polivios, my friends from my University years in Greece. Thank you to all the people in graphics that I met in conferences – their enthusiasm kept me going and their ideas inspired me!

This work would never be possible without the support of my family, my dad, mum and brother. This thesis is dedicated to them.

To my parents and brother

‘ In general, the impossible must be justified by reference to artistic requirements, or to the higher reality, or to received opinion. With respects to the requirements of art, a probable impossibility is to be preferred to a thing improbable, and yet possible. Again, it may be impossible that there should be men such as Zeuxis painted. ‘Yes’ we say, but the impossible is the higher thing; for the ideal type must surpass the reality. To justify the irrational, we appeal to what is commonly said to be. In addition to which, we urge that the irrational sometimes does not violate reason; just as it is probable that a thing may happen contrary to probability.’

Aristotle, Poetics, Book XXV, Paragraph 16-18

‘Γενικά, το αδύνατο πρέπει να δικαιολογεί κανείς αναφερόμενος ή στις απαιτήσεις της ποίησης ή στην εξιδανείκευση ή στην κοινή πίστη. Γιατί για την ποίηση προτιμότερο είναι το πιθανό αδύνατο παρά το απίθανο και δυνατό. Αφ’ ετέρου, ίσως είναι αδύνατο να υπάρχουν άντρες όπως ζωγράφισε ο Ζεύξης, αλλά είναι ιδανικότερο τούτο διότι το υπόδειγμα πρέπει να υπερέχει του πραγματικού. Αναφερόμενος σε ότι λέγουν πρέπει να δικαιολογεί τα παράλογα με αυτό τον τρόπο. Κατά καιρούς, τα παράλογα δεν είναι απίθανα, διότι είναι πιθανό μερικά, αν και απίθανα, να γίνονται.’

Αριστοτέλης, Περί Ποιητικής, Βιβλίο XXV, Παράγραφος 16-18

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One aim of interactive computer graphics and Virtual Environment (VE) technologies for simulation and training is the creation of accurate, high-quality imagery which faithfully represents a physical environment. The ultimate goal, as often argued, is to create synthetic spaces which are going to induce a sense of 'presence' similar to the real world. The accurate simulation of real-world spaces and illumination could be required as well as the simulation of interaction interfaces to match the human perceptual and motor systems. Reliable usability evaluation techniques and robust metrics are essential in order to assess VE implementations comprising of computer graphics imagery, display technologies and 3D interaction metaphors across a range of application fields. One common belief is that efficient task performance measures should serve as fidelity metrics for any application that requires a high level of simulation fidelity and mainly targets, for instance, transfer of training in the real world. A commonly employed strategy, therefore, for assessing the simulation fidelity of a Virtual Environment is to compare task performance in that environment in relation to task performance in the real world scene represented in the VE. Studies, though, that focus on task performance efficiency are often limited to a specific application and are not based on specific, formally designed frameworks but in most cases on arbitrary selected tasks. Without a formal framework, relevant studies could reveal contradictory results.

Another common approach is to take a cross-application construct, such as the sense of 'presence' and assess the effectiveness of a VE, or aspects of a VE according to its success in enhancing presence. There is a widespread belief that presence should somehow improve task performance, although this has never been verified or indeed reasons given as to why this should be the case.

This thesis introduces a new approach. It argues that because of the wide-ranging VE applications and differences in human users across their backgrounds, abilities, methods of processing information and susceptibility to aftereffects, attempts to validate VEs on the basis of task performance alone are unlikely to succeed. The thesis uses a methodology that attempts to understand *how* tasks are undertaken within a VE rather than *what* is achieved. Hence, task performance is not the only issue. The cognitive strategies participants employ to complete a task and the relationship between these and the technology employed in specific applications need to be examined.

Spatial perception or spatial memory is a task often incorporated in benchmarking processes, since spatial awareness is crucial for human performance efficiency. The utility of VEs for any applications for which they are being proposed is predicated upon the accuracy of the spatial representation formed in the VE which could be tested employing a memory task after exposure to a VE implementation. For this reason spatial awareness is going to be the focus of this thesis. A unified theory of spatial perception is presented which will focus on elements of human memory other than the just the amount of accurate memory recall, in an effort to unravel the actual mental strategies that humans employ in order to form mental models of a space. A general framework as such is going to offer a wealth of information about human perception related to rendering algorithms, sophisticated displays such as Head Mounted Displays (HMDs) and interaction interfaces such as the common mouse or head tracking.

The framework presented in this thesis is drawn from traditional memory research. This framework is adjusted and applied to form an experimental procedure for acquiring human judgements of memory recall and memory awareness states in real scenes and their computer graphics simulation counterparts. This theory, towards assessing simulation fidelity of a computer graphics implementation, is not only focusing on accurate spatial awareness and memory recall but also on the mental processes or awareness states humans employ in order to complete a memory task [Tul85], [Tul93], [KG94], [Gar2001]. Participants could describe how they make their spatial recollections after exposure to a VE application by selecting between four choices of awareness states ('remember', 'know', 'familiar' and 'guess'). These depend on the level of visual mental imagery involved during retrieval, the familiarity of the recollection and also include guesses, even if informed. In order to demonstrate the differences between using task performance based metrics and human evaluation of cognitive awareness states, a set of VEs displayed mainly on Head Mounted Displays (HMDs) were created. Resulting scenes were then compared to the real scene they represented by employing the spatial memory awareness states methodology as well as assessments of presence, simulator sickness and responses to lighting. By employing methodologies that have been examined and validated through decades of experimentation such as the memory awareness states methodology employed in this thesis, computer graphics research gets closer to successfully exploiting the human perceptual mechanisms.

1.1 Contributions

- The literature review is wide-ranging. It covers all the different aspects of this thesis including computer graphics rendering, photometry and radiometry specifics, image quality metrics, fundamental literature on VE technologies and displays including task performance issues, presence and simulator sickness (aftereffects), human factors approaches and human memory research.
- A VE experimental framework is presented that assesses the simulation fidelity of a VE application based on a validated memory model that focuses on the cognitive processes or memory awareness states of participants completing a memory task. The statistical analysis is based on prior and posterior probabilities.

- Three main studies, which employ the memory awareness states methodology, form the core of the work in this thesis involving more than 200 participants. The technology involved in these experiments includes Head Mounted Displays, head-tracking interfaces and stereo computer graphics imagery.
- Experimental data are analysed by calculating prior and posterior probabilities. Relevant results showed that although task performance (accurate memory recall) did not differ between conditions such as the HMD-related conditions, nevertheless participants adopted different strategies of recall. This indicates that the simulation of task performance does not necessarily lead to simulation of the cognitive strategies humans employ to achieve tasks. Certain cognitive strategies could be desirable for specific applications. Also, head tracking as an interaction interface proved to induce a higher amount of accurate recollections retained in time in comparison to mouse interfaces, although there was a shift of cognitive processes associated with these scores.

Generally, such a detailed insight, shown by results mentioned above into spatial perception in a computer graphics world could not be possible if the commonly employed task performance-based techniques are adopted.

1.2 Thesis Outline

Chapter 2: Background

Chapter 2 initially introduces a set of fundamental terms for computer graphics rendering, starting with defining light and its properties, light energy, photometry, radiometry and illumination models. This section concludes with an overview of visual perception principles relevant to computer graphics followed by an overview of research on perceptually-based rendering (employing knowledge of visual perception for the benefit of computer graphics) as well as on perceptually-based image quality metrics. The main body of this chapter focuses on interactive computer graphics scenes, presenting an overview of human factors issues such as human performance efficiency and aftereffects related to VE technologies and displays. The notion of presence and its relation to task performance and aftereffects is analysed concluding with a presentation of literature relevant to comparisons of real scenes with their computer graphics simulations.

Chapter 3: A Methodology based on Memory Semantics

Chapter 3 introduces the methodology of spatial awareness states by initially offering an overview of related theories drawn from classical memory research. Next, relevant spatial perception literature relevant to VEs is critically reviewed. Finally, a spatial perception methodology is described. This results in a system that enables the comparison of synthetic imagery to the real environment being represented by employing spatial tasks under varied display and interface conditions, using human participants.

Chapter 4: A Preliminary Study

The framework introduced in Chapter 3 is applied in this chapter in a preliminary study informally designed comparing a real task situation with its computer graphics simulation counterpart, demonstrating the feasibility and applicability of the memory awareness states approach. The task employed is a memory recall task and the conditions include the real world, desktop monitor, HMD and audio-only.

Chapter 5: Main Experiments and Results

Building on the preliminary study in this chapter, a formally designed set of experiments is presented incorporating the framework presented in Chapter 3. A focused visual spatial memory task is employed incorporating accurate photorealistic monocular and stereo rendering based on photometry measurements and more complex interaction interfaces such as head tracking. The conditions include the real world, desktop monitor focusing on immersive Head Mounted Display implementations. The main study is repeated after a week across all conditions. The results are analysed in detail together with a discussion on how these results could provide an aid towards validating a VE application with respect to the real scene it represents as well as improve the efficiency of assembling a VE implementation for a specific application.

Chapter 6: Conclusions

Finally, the results and contributions of this thesis are presented. Future work, unveiled by relevant conclusions is suggested.

Chapter 2

Background

Increased applications of real-time, interactive computer graphics mainly for simulation and training has made it crucial to examine the way these implementations are evaluated. Relevant applications could be displayed on typical desktop monitors or onto more sophisticated *immersive* visual displays such as Head Mounted Displays (HMDs). While the advances of display technologies [Br99] are quite breathtaking, what is still lacking is a set of robust metrics with which to assess the quality of visual imagery, interfaces and overall designs employed for Virtual Environment (VE) applications [BZSS95]. In this chapter, related research in the fields of computer graphics, visual perception and usability engineering for virtual environment technologies is reviewed separately. The first section of this chapter describes fundamental concepts for computer graphics rendering of 2D scenes, such as light properties and illumination models employed in this thesis focusing on perceptually-based rendering, image quality metrics and comparison studies of real world scenes with their 2D computer graphics simulations counterparts. Relevant research incorporates principles of human visual perception to improve computer graphics rendering and develops metrics that assess the quality of computer graphics scenes.

The second section proceeds to analyse concepts relevant to interactive 3D computer graphics worlds and VE technologies focusing on human factors issues, the notion of presence and comparison studies of real world scenes with interactive 3D computer graphics simulations incorporating sophisticated display technologies.

2.1 Computer Graphics Rendering

What is a *realistic* image, whether painted, photographed or computer-generated? Figure 2.1 shows the most famous painting by Magritte in which the distinction between illusion and reality is called into question.

Figure 2.1: ‘The Human Condition’ by Magritte, 1933, Oil on canvas, National Gallery of Art Washington DC.

The term ‘*realistic*’ is used broadly to refer to an image that captures and displays the effects of light interacting with physical objects. Thus, realistic images are treated as a continuum and the techniques used to create them are characterised as ‘more’ or ‘less’ realistic. Images referred to as *'photorealistic'* attempt to synthesise the field of light intensities that would be focused on the film plane of a camera aimed at the objects depicted [FDFH90].

The quest for visual realism for static (2D) or interactive (3D) environments is a challenge for the field of computer graphics. The production (rendering) of realistic imagery requires a precise generation of lighting effects, which involves the simulation of physical phenomena including light emission, propagation and reflection as well as the simulation of physical entities as objects and space. Rendering systems can now approximate the physical distribution of light in an environment. However, physical accuracy of light and geometry does not guarantee that the displayed images will have authentic appearance to the human eye or that they will seem 'real'.

2.1.1 The Physical Behaviour of Light

Light is one form of *electromagnetic radiation,* a mode of propagation of energy through space that includes radio waves, radiant heat, gamma rays and X-rays. One way in which the nature of electromagnetic radiation can be pictured is as a pattern of waves propagated through an imaginary medium. The term 'visible light' is used to describe the subset of the spectrum of electromagnetic energy to which the human eye is sensitive. This subset, usually referred to as the *visual range* or the *visual band* consists of electromagnetic energy with wavelengths in the range of 380 to 780 nanometres, although the human eye has very low sensitivity to a wider range of wavelengths, including the infrared and ultraviolet ranges. The range of visible light is shown in Figure 2.2. As shown, the wavelength at which the human eye is most sensitive is 555 nm.

In the field of computer graphics three types of light interaction are primarily considered: *absorption, reflection and transmission*. In the case of absorption, an incident photon is removed from the simulation with no further contribution to the illumination within the environment. Reflection considers incident light that is propagated from a surface back into the scene and transmission describes light that travels through the material upon which it is incident and can then return to the environment, often from another surface of the same physical object. Both reflection and transmission can be subdivided into three main types:

Specular: When the incident light is propagated without scattering as if reflected from a mirror or transmitted through glass.
Diffuse: When incident light is scattered in all directions.
Glossy: This is a weighted combination of diffuse and specular.

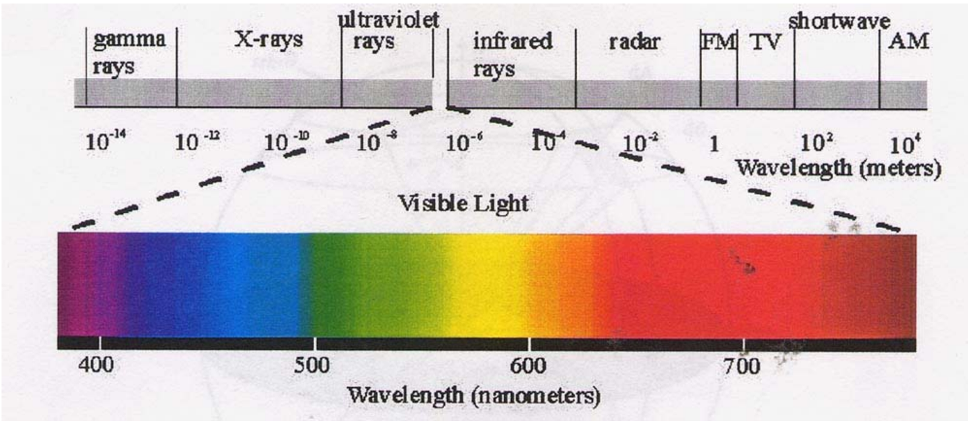


Figure 2.2: The visible portion of the electromagnetic spectrum.

Most materials do not fall exactly into one of the material categories described above but instead exhibit a combination of specular and diffuse characteristics.

2.1.2 Radiometry and Photometry

The field of lighting simulation is concerned with the accurate physical modelling of light propagation through an environment. The simulation of light should employ measures in units that could allow comparisons of real life scenes with its simulation counterpart. Two related methods exist for the measuring of light distributions in simulations [Kaj90]:

- **Radiometry** is the science of measuring radiant energy from any part of the electromagnetic spectrum using optical instruments which mainly measure light in the visible, infrared and ultraviolet wavelength. Radiometry is used in graphics to provide the basis for illumination calculations. Relevant metrics for computer graphics are the following:

Radiant Energy, measured in *Joules*: Light is radiant energy.
Radiant Flux, measured in *Watts*: The radiant energy flowing through an area per unit time.
Radiant Flux Density measured in *Watts per square metre*: The quotient of the radiant flux incident on or emitted by surface element surrounding the point and area of the element.
Radiant Exitance (Radiosity), measured in *Watts per square metre*: The radiant flux leaving the surface per unit area of the surface.
Irradiance, measured in Watts per square metre: The radiant flux incident on the receiver per unit area of the receiver.
Radiant Intensity, measured in *watts per steradian*: The radiant flow from a point source in a particular direction.
Radiance measured in *watts per steradian per metre squared*: Radiant flux arriving or leaving from a surface, per unit solid angle per unit projected area.

- **Photometry** is the science of measuring light within the visible portion of the electromagnetic spectrum in units weighted in accordance with the sensitivity of the Human Visual System (HVS). Photometry's goal is to measure the subjective impression produced by stimulating the human-visual system with radiant energy. Photometry is just like radiometry except that everything is weighted by the spectral response of the eye. Visual photometry uses the eye as a comparison detector, while physical photometry uses either optical radiation detectors constructed to mimic the spectral response of the eye, or spectroradiometry coupled with appropriate calculations to do the eye response weighting. Within the visual range, the sensitivity of the eye varies according to wavelength as shown in Figure 2.3. The photopic metrics relevant to computer graphics are the following:

Light: Radiant energy that is capable of exciting the retina and producing a visual sensation (visible energy).

Luminous Flux, measured in *Lumens*: The rate of flow of light in respect to time. Since 'light' is visible energy, the lumen refers only to visible power.
Luminous Factor, measure in *Lumen/Watt*: The ratio of the luminous flux at that wavelength to the corresponding radiant flux. It expresses the sensitivity of the human eye to the visible wavelengths.
Luminous Intensity measured in *Candelas*: The luminous flux per solid angle emitted or reflected from a point.
Illuminance, measured in *Lumen per square metre (Lux)*: The area density of the luminous flux incident on a surface.
Luminous Exitance: The total area density of luminous flux leaving a surface at a point.
Luminance, measured in *Candelas per square metre*: The area density of the luminous flux incident on a surface in a specific direction. This is the measurable quantity closest to brightness. Brightness is a subjective attribute to light to which humans perceive as being very dim or very bright.

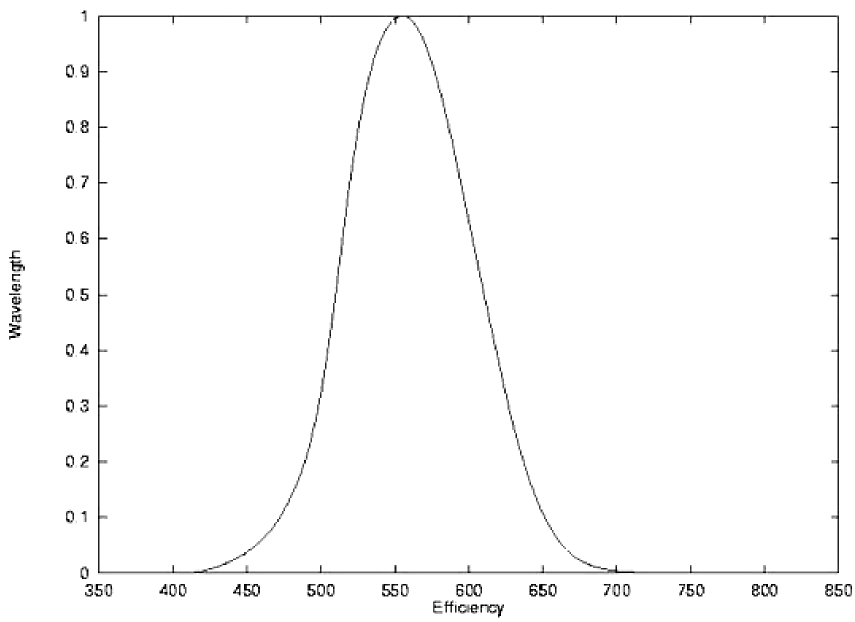


Figure 2.3: Luminous Efficiency Curve.

In this thesis, photometric measurements are going to be acquired from a real-world space to ensure accuracy of illumination in the computer graphics simulation of that space which is, subsequently, employed for experimentation (Chapter 5).

2.1.3 Computer Graphics Illumination Models

An illumination model computes the colour at a point in terms of light directly emitted by the light source(s) [FDFH90]. A *local illumination model* calculates the distribution of light that comes directly from the light source(s). A *global illumination* model additionally calculates reflected light from all the surfaces in a scene which could receive light indirectly via intereflections from other surfaces. Global illumination models include, therefore, all the light interaction in a scene, allowing for soft shadows and colour bleeding that contribute towards a more photorealistic image. The rendering equation expresses the light being transferred from one point to another [Kaj86]. Most illumination computations are approximate solutions of the rendering equation:

$$I(x,y) = g(x,y)\left[\,\varepsilon(x,y) + \int_S p(x,y,z)\, I(y,z)\, dz\,\right]$$

where

x,y,z are points in the environment,

$I(x,y)$ is related to the intensity passing from y to x,

$g(x,y)$ is a 'geometry' term that is 0 when x,y are occluded from each other and 1 otherwise,

$p(x,y,z)$ is related to the intensity of light reflected from z to x from the surface at y,

the integral is over all points on all surfaces S.

$\varepsilon(x,y)$ is related to the intensity of light that is emitted from y to x.

Thus, the rendering equation states that the light from y that reaches x consists of light emitted by y itself and light scattered by y to x from all other surfaces which themselves emit light and recursively scatter light from other surfaces. The distinction between *view-dependent* rendering algorithms and *view-independent* algorithms is a significant one. *View-dependent* algorithms discretise the view plane to determine points at which to evaluate the illumination equation, given the viewer's direction, such as ray-tracing [Gla95]. *View-independent* algorithms discretise the environment and process it in order to provide enough information to evaluate the illumination equation at any point and from any viewing direction, such as radiosity [CG85].

Radiosity algorithms display view-independent diffuse interreflections in a scene. These algorithms assume the conservation of light energy in a closed environment. All energy emitted or reflected by every surface is accounted for by its reflection from or absorption by other surfaces. The rate at which energy leaves a surface is the sum of the rates at which the surface emits energy and reflects or transmits it from that surface or other surfaces. Radiosity methods allow any surface to emit light; thus, all light sources are modelled inherently as having area. The surfaces of a scene are broken up into a finite number of *n* discrete patches, each of which is assumed to be of finite size, emitting and reflecting light uniformly over its entire area, as shown in Figure 2.4. Thus, if each patch is considered to be an opaque diffuse emitter and reflector, then for surface *i*,

$$B_i = E_i + p_i \sum_{1 \le j \le n} Bj\ Fji \frac{Aj}{Ai}$$

where B_i and B_j are the radiosities of patches *i* and *j*,
E_i is the rate at which light is emitted from patch *i* ,
p_i is patch *i*'s reflectivity,
F_{ji} is the form factor which specifies the fraction of energy leaving patch *j* and arrives at patch *i, j!=i.*

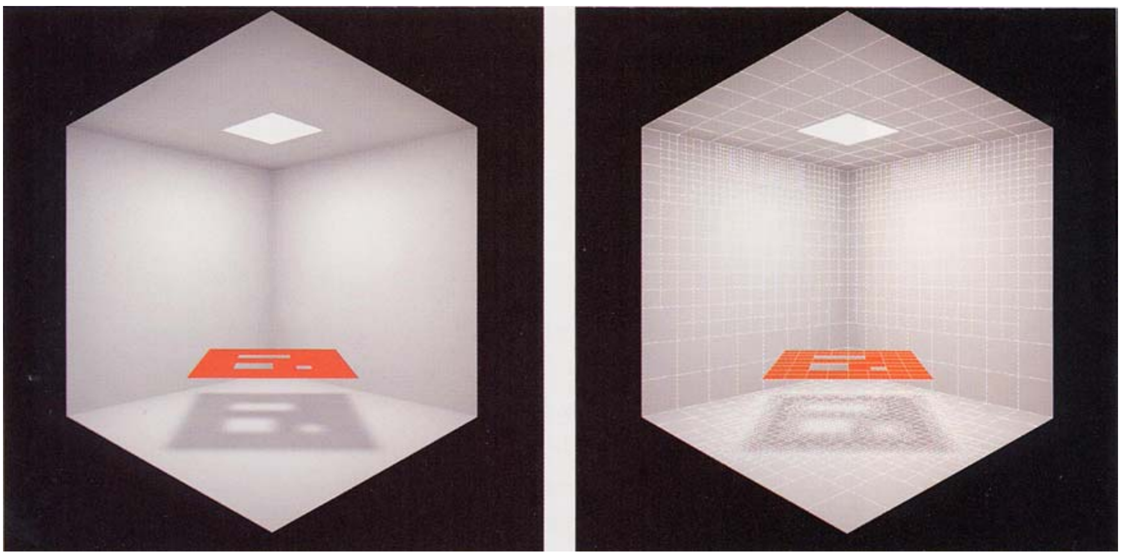


Figure 2.4: Radiosity Example. The area in shadow still receives some illumination. Image by Kim Wagner Jensen, copyright Hewlett Packard company.

The radiosity equation states that the energy leaving a unit area of surface is the sum of the light emitted (from within a volume itself) plus the light reflected (sum of all the rays coming to that point). The *progressive refinement* radiosity algorithm allows partial solutions to be displayed early on in the computation, although perhaps inaccurately, which can be successively refined to greater accuracy as more computing time is allocated. Generally, the result of a radiosity solution is not just a static image but an interactive three-dimensional representation of light energy in an environment (Figure 2.5). The radiosity algorithm is going to be employed to render the computer graphics interactive scenes in the Chapter 5 of this thesis.

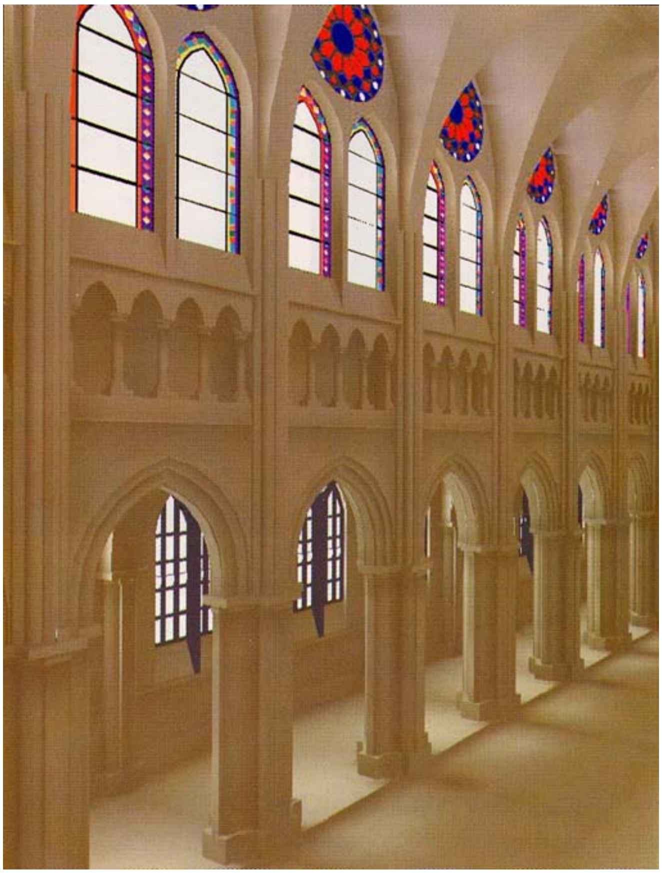


Figure 2.5: Synthetic radiosity reproduction of the cathedral in Chartres, France. Image by John Wallace and John Lin, copyright Hewlett Packard Company.

2.2 Visual Perception in Realistic Image Synthesis

In recent years, research in realistic image synthesis has included perceptually based rendering, considering aspects of the Human Visual System (HVS) to produce faster or more realistic computer graphics images. The basic goal of realistic rendering is to create images perceptually indistinguishable from real scenes. Since the human observer judges the fidelity and quality of the resulting images, the perceivable differences between the appearance of a computer graphics image and its real world counterpart should be minimised. Thus, visual perception issues are clearly involved in realistic rendering and should be considered at various stages of computation, rendering and displaying [CMTDM2000]. This specific research direction has gained much attention of the computer graphics research community [Gr99], motivated by the progress in physiology, psychophysics and psychology in providing computational models of the HVS [BGG96]. In this section, principles of human visual perception that are employed in this thesis (mainly in Chapter 4 and 5) are going to be reviewed. Also, research that utilises principles of human visual perception towards efficient computer graphics algorithms and image quality metrics is going to be mentioned. It is useful to demonstrate that such knowledge is invaluable for the progress of the computer graphics field.

2.2.1 Human Visual Perception

The ability of a person or animal to detect fine spatial pattern, therefore, resolve detail in an image, is expressed as *visual acuity* [BGG96]. The human eye is more sensitive to intermediate changes in brightness as opposed to gradual or sudden changes. Acuity decreases with increase in distance. At any instant, the human eye samples a relatively large segment of the optic array (the *peripheral* field) with low acuity and a much smaller segment (the *central*, or *foveal* field) with high acuity. Saccadic movements are rapid jumps of eye position in order to focus to an object. Smooth and saccadic eye movements shift this high-acuity segment about rapidly so that acute vision over a wide angle is achieved. If the distance of an object from the observer changes, *convergence* movements keep it fixated by the foveal field of both eyes. The Human Field of Vision (FoV) is normally around 200 degrees.

Perceptual constancy refers to the fact that we perceive a surface as having a constant appearance despite changes in the spectral composition of light reflected from it. A number of perceptual constancies have been identified:

- *Colour constancy* refers to the fact that the Human Visual System perceives a surface as having a constant colour despite changes in the type of illumination, e.g. spectral distribution of illumination. An everyday example of colour constancy occurs when we move from daylight to an artificially lit room. The illumination from a common electric bulb is different from that of sunlight and is relatively richer in long wavelengths. The same surface will therefore reflect more long wavelength light under a light bulb than in sunlight, and, if our perception of colour depended on wavelength alone, it would appear redder. In fact, the colour actually perceived remains largely constant; a white sheet of paper, for example, does not appear redder indoors and bluish outdoors. Colour constancy is not always perfect, depending on the individual but generally does not vary a lot, however, those with a 'good' eye for colour account for this when, for example, shopping for clothes under artificial light [BGG96].
- *Lightness constancy* refers to the ability of the HVS to perceive surface lightness as constant despite the changes in illumination.
- *Shape constancy* refers to the ability of the HVS to perceive objects as having the same shape despite changes in their orientation. A number of researchers suggest that our usual ability to recognise objects across a range of viewpoints arises as a result of our experiencing and storing different viewpoints separately [BE92], or through the recognition of viewpoint-invariant features [Bie87b].

Depth perception is the ability to see the world in three dimensions. *Binocular disparity* (horizontal disparity is the distance between the two views of the two eyes) and *pictorial cues* provide depth perception information. Animals with overlapping visual fields have stereoscopic information available to them from a comparison of the images obtained at the two eyes. Each eye sees a slightly different view of the world due to the horizontal distance between the two eyes. Objects at different distances will appear to move together, or apart, reflecting the horizontal disparity between the two views. The human brain fuses the two separate views into one that is interpreted as being in 3D. The two images are called a *stereo pair*. Figure 2.6 shows the geometry of binocular vision. Most theories of stereo vision have argued that in analysing binocular disparity the visual system must determine which parts of one eye's image correspond to particular parts in the other eye's image.

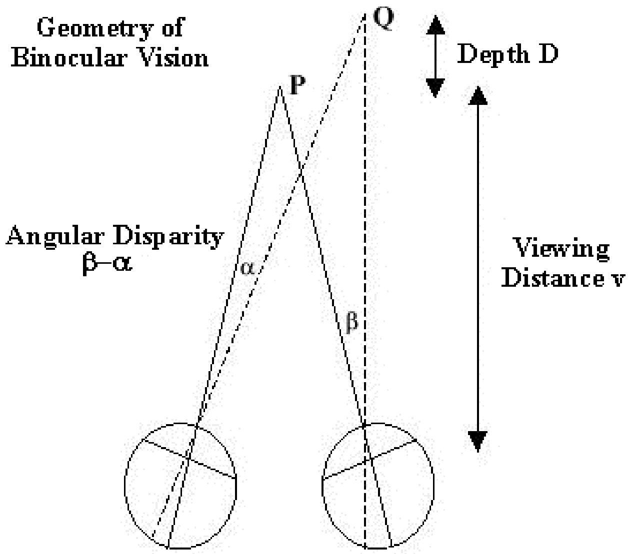


Figure 2.6: Geometry of binocular vision.

In Figure 2.6, the eyes fixate at point P. Point Q further away than P is imaged on non-corresponding or disparate points in the two eyes. The disparity ($\beta-\alpha$) produced by point Q is measured by the discrepancy between the two image locations in degrees of visual angle. Disparity is proportional to depth, D, but inversely proportional to the squared viewing distance [BGG96]. For instance, if the image of Q was located 6 degrees away from the fovea in one eye but 5 degrees away in the other, then Q produced an angular disparity of 1 degree. Disparity increases with the amount of depth but decreases rapidly with increasing viewing distance.

In computer graphics/VE displays various techniques exist for providing different images to each eye including glasses with polarised filters and holography. Some of these techniques make possible true 3D images that occupy space, rather than being projected on a single plane. These displays can provide an additional depth cue: Closer objects actually are closer just as in real life so the viewer's eyes focus differently on different objects, depending on each object's proximity.

On a 2D display, stereopsis can be achieved by computing separate images for the left and right eyes and channelling each image to the respective eye. The three most common ways of achieving stereo displays are:

- *Dual Channel* which is also referred to as separate channel or multi-channel, this form of stereo uses two separate video signals and it is typically used with Head Mounted Displays (HMDs).
- *Line Interleaved* which is also referred to as Interlaced or Passive Stereo, this form uses polarised glasses (usually inexpensive passive glasses) to isolate left and right eye scan lines.
- *Quad Buffered* which is also referred to as Field Sequential, this form uses shuttering glasses to sequentially produce alternating left and right eye images at a rate higher than the human brain is able to distinguish (usually expensive mechanical shutter glasses).

Pictorial cues to depth were employed by artists since the Renaissance to convey a sense of depth in their work. Many of these pictorial cues are variations of *perspective*. They arise from the way in which a 3-D world is projected onto a 2-D retina, from a particular viewpoint. Linear perspective is perhaps the best known pictorial cue to depth. Other cues for depth include texture gradient, shading, height and motion perspective or parallax.

Depth perception may also be involved in the effect of **monocular foreground occlusions** on the perception of static, two-dimensional (2D) pictures. The literature on monocular foreground occlusions for viewing static, 2D pictures is reviewed by Rogers [Rog95]:

> ‘It has long been observed that the appearance of three dimensions in a picture is more striking under certain viewing conditions, such as viewing the picture with one eye only and by using a peephole or a lens in a reduction screen or viewbox. Numbers of old viewing devices make use of one method or another to produce their sometimes powerful effects. The usual explanation is that the restricted view enhances the effectiveness of pictorial–depth information by reducing the conflicting flatness information that specifies a picture’s objective surface. Peephole viewing is monocular and head motion is

prevented, and this should enhance perceived depth. A peephole itself also restricts the observer's view of the picture itself hiding the frame and surrounding surfaces. Loss of the visible frame and discontinuous surrounding surfaces reduce information for the picture as a flat object (perhaps even for the presence of a surface at all), potentially enhancing the illusion of depth in the picture. This is the oft-cited reason for the success of the various picture-viewing devices.'

To actually experience the foreground occlusion effect, a simple technique is to hold up a cardboard box with a hole in it which just blocks the edges of a monitor. For best results, one could play on the monitor a scene with considerable movement. Compare one's impressions looking through the hole with one eye to an unobstructed view of the monitor at the same distance, again with only one eye [Pro98].

2.2.2 Perceptually Driven Rendering

Since global illumination solutions are costly in terms of computation, there are good prospects for their efficiency improvement by focusing computation on those scene features which can be readily perceived by the human observer under given viewing conditions. The features that are below perceptual visibility thresholds can be simply omitted from the computation without causing any perceivable difference in the final image appearance [CMTDM2000]. For instance, exploiting the poor colour spatial acuity of the HVS, Meyer and Liu [ML92] developed an adaptive image synthesis algorithm that uses an opponents processing model of colour vision comprising chromatic and achromatic colour channels. They achieved a modest saving in computational effort and showed, using a psychophysical experiment that decreasing the number of rays used to produce the chromatic channels had less of an effect on image quality than reducing the number of rays used to create the achromatic channels. This was the first work to attempt to minimise the computation of colour calculations. Subsequently, the research community has incorporated more complex models of the HVS into the rendering algorithms [Mys98b], [BM98b].

Another technique that addresses the problem of mapping real-world illumination to a display is *tone mapping* which takes advantage of the HVS sensitivity to relative luminance rather than absolute luminance. If the luminance of a light source is increased 10 times, the human viewer does not perceive that the actual brightness has increased 10 times. The actual relationship is logarithmic meaning that the sensitivity of the eye decreases rapidly as the luminance of the source increases. Tumblin & Rushmeier [TR93] were the first who attempted to match the brightness of a real scene (vast luminances and contrast ratio) to the brightness of the computer graphics image on a typical display (limited display luminance and contrast ratio). They attempted to recreate the same perceptual response for a human viewer looking at a display as the response in the real world. For a complete review, see [CMTDM2000].

Generally, using validated visual models that could predict image fidelity, researchers can work towards greater efficiency and speed in rendering, knowing that resulting images will still be faithful visual representations to the human eye [Mcn2000].

2.2.3 Perceptually based Image Quality Metrics

Current global illumination algorithms usually rely on energy-based metrics of solution errors, which do not necessarily correspond to the *visible* improvements of the image quality [LSG94]. Ideally, the development of perceptually-based error metrics which can control the accuracy of every light interaction between surfaces, are desired. This can be done by predicting the visual impact those errors may have on the perceived fidelity of the rendered images. Another approach is to develop a perceptual metric which operates directly on the rendered images. If the goal of rendering is just a still frame, then the image-based error is adequate. In the case of view-independent solutions, the application of the metric becomes more complex because a number of ‘representative’ views should be chosen. In practice, instead of measuring the image quality in absolute terms, it is much easier to derive a relative metric that predicts the perceived differences between a pair of images [RGPSR95].

It is commonly known that a common mean squared-error metric usually fails in such a task. An example of an advanced image fidelity metric that incorporates a complex HVS model is the Visible Differences Predictor (VDP) [Dal93]. Based on previous human experimentation, the VDP predics many characteristics of human perception. The VDP metric when applied in global illumination computation provides a summary of the algorithm performance as a whole rather than giving a detailed insight into the work of its particular elements. The VDP gets as input a pair of images and it generates as output a map of probability values that characterise if these differences could be perceived by a human observer. Another perceptually based metric is the Sarnoff Visual Discrimination Model that focuses on modelling the physiology of the visual pathway [Lub95].

2.2.4 Comparing Real and 2D Computer Graphics Scenes

As the goal of realistic image synthesis is to generate accurate representations of real-world scenes, synthetic images should be compared to their real counterparts using human observers and taking into account aspects of the HVS. Since this is the general methodology employed in this thesis, research that focuses on comparisons between real-life and computer graphics simulations is going to be reviewed. In this section, research that attempts to compare a real-life scene with a computer graphics static scene is presented. Later in this chapter (section 2.5.4), research focusing on comparisons of interactive computer graphics scenes with real-world task situations, displayed on more sophisticated displays is going to be reviewed.

The first research effort to compare real and simulated **static** scenes side by side was attempted by Meyer et al. [MRCG86]. They used a 5-sided cube as their test environment, which was placed in a dark room (Cornell box). Radiometric values predicted using a radiosity rendering of a basic scene were compared to physical measurements of radiant flux densities in the real scene. The results of the radiosity calculations were transformed to the RGB values from display following the principles of colour science. Measurements of irradiation were made at 25 locations in the plane of the open face for comparison with the simulations. Meyer et al. then proceeded by transforming the validated simulated scene to values displayable on a television monitor. Twenty participants were asked to differentiate between a real environment and the displayed image, both of which were viewed through the back of a view camera. They were asked which of the images was the real scene. 45% of the participants indicated that the simulated image was actually the real scene, a result close to chance, e.g. overall random selections. Participants considered the overall match and colour match to be good, however, some weaknesses were noticed in the sharpness of the shadows, mostly due to the brightness of the ceiling panel, caused by the orientation of the light source. Although there is strong support for the perceptual validity of the simulation and display process in this study, there were also some shortcomings: The scene was very simple and the results showed that the participants could have just guessed.

Another approach towards comparing real and simulated static scenes takes a captured image of the real scene in question and uses numerical techniques to determine the perceptual differences between the two [RGPSR95]. Rushmeier et al. introduced components of a perceptually based metric using ideas from the image compression literature. Image compression techniques seek to minimise storage space by saving only what will be visible in an image. The goal of this study was to obtain results from comparing two images using these models that were large if large differences between the images exist and small when they are almost the same. The model inspired by Daly [Dal93] had good results and was also the only one that considers human limits in dark adaptation. It was noted in this study that the biggest challenge is to take insights into human perception and apply them to visual simulation directly, computing only as much as is needed to satisfy the observer. Translating a perceptual model into a progressive calculation is not a straightforward task, however, this is not the focus of this thesis.

Myszkowski [Mys98b] completed a comprehensive validation and calibration of the Visible Differences Predictor (VDP) response [Dal93], using psychophysical experiments. He subsequently used the VDP local error metric to steer decision making in adaptive mesh subdivision and in isolating regions of interest for more intensive global illumination computations. The VDP was tested to determine how close VDP predictions come to subjective reports of visible differences between images by designing two human psychophysical experiments. Results from these experiments showed a good correspondence between human observations and VDP results. Figure 2.7 shows a series of psychophysical experiments conducted by Drago & Myszkowski [DM2001]. Participants were asked to rate the degree of realism (Figure 2.7, right) between two computer graphics renderings and a photograph of a real scene (Figure 2.7, left); all of them representing the atrium of the University of Aizu in Japan.

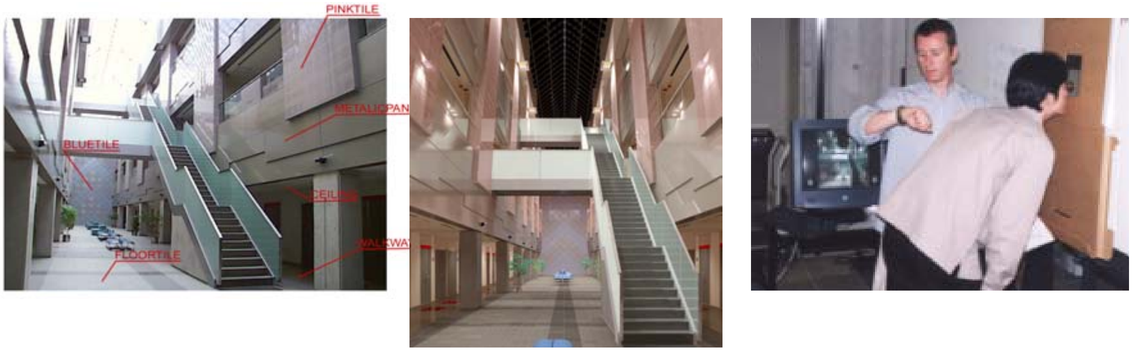


Figure 2.7: Subjective responses of perceived realism [DM2001].

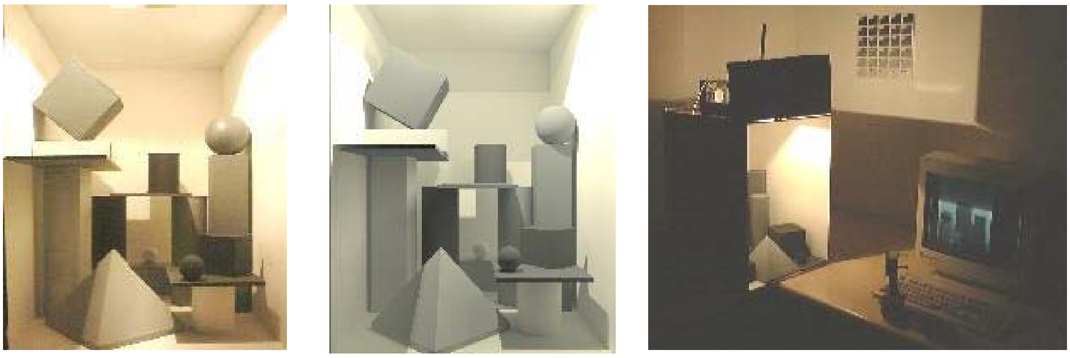


Figure 2.8: Comparing real and synthetic scenes using human judgements of lightness perception [MCTG2000].

One of the renderings was rendered accurately in terms of the illumination, however, artists were called to contribute towards the second rendering according to their aesthetic recommendations. Results showed that the rendering of accurate illumination was ranked higher (in terms of the perceived degree of realism) in relation to the rendering artistically composed (Figure 2.7, middle).

In a more recent approach, McNamara et al. [MCTG2000] introduced a method for measuring the perceptual equivalence between a real scene and a computer simulation of the same scene. The model developed in this research is based on psychophysical experiments; more specifically on human judgements of lightness when viewing a real scene (Figure 2.8, right), a photograph of the real scene (Figure 2.8, left) and nine different computer graphics simulations including a poorly meshed radiosity solution and a raytraced image. Results were produced through a study of vision from a human rather than a machine vision point of view. They showed that certain rendering solutions, as the tone-mapped one, were of the same perceptual quality as a photograph of the real scene (Figure 2.8).

2.3 Interactive Computer Graphics Scenes

According to Barfield and Furness [BF95], a *Virtual Environment (VE)* is defined as a representation of a computer model or database that can be *interactively* experienced and manipulated by the VE participants. Ellis [El95] characterised a Virtual Environment as a 'synthetic, interactive, illusory environment perceived when a user wears or inhabits appropriate apparatus providing a co-ordinated presentation of sensory information mimicking that of a physical environment'. Also, in [El94] Virtual Environments are defined as interactive *virtual image* displays enhanced by non-visual display modalities, such as auditory and haptic, to convince users that they are immersed in a synthetic space. A virtual image is defined as the visual, auditory, tactile and kinaesthetic stimuli that are communicated to the sensory receivers of the participants such that they seem to originate from within the three-dimensional space [BF95].

An interface is a communication medium between a human and the functional elements of a machine. A *virtual interface* is a system of signals coming from software or hardware that form an interactive medium through which, firstly, information is communicated to the user's senses in the form of three-dimensional images and possibly tactile and kinaesthetic feedback or sound, and secondly, participants' actions are monitored manipulating the virtual environment [BF95].

2.3.1 Origins of Virtual Environments

Experiencing imagined environments has been rooted in the human mind originating as early as human cave art. Lewis Carroll's *'Through the Looking Glass'* describing the adventures of Alice in the Wonderland, is a more modern example of this fascination. An environment which would allow its 'inhabitants' to move about and manually interact with computer graphics objects was envisioned in the science fiction plot of William Gibson's '*Neuromancer*' in 1984. Yet, Ivan Sutherland made the first actual implementation of such a system using a head-mounted stereo display possible much earlier, in the middle 1960's. He developed fast graphics hardware in 1965 (Figure 2.9), specifically for experiencing computer-synthesised environments through head-mounted graphics displays [Su65], [Su70]:

> 'Don't think of that thing as a *screen*, think of it as a *window*, a window through which one looks into a virtual world. The challenge to computer graphics is to make that virtual world look real, sound real, move and respond to interaction in real-time and even feel real.'

Myron Krueger developed another early implementation of a synthetic, interactive environment in the 1970's [Kru77], [Kru85]. Unlike the device developed by Sutherland, this environment called VIDEOPLACE was projected onto a wall-sized screen. When a participant touched an object on the screen, the system generated a graphics or auditory response. This system did not limit participants' positioning or movement.

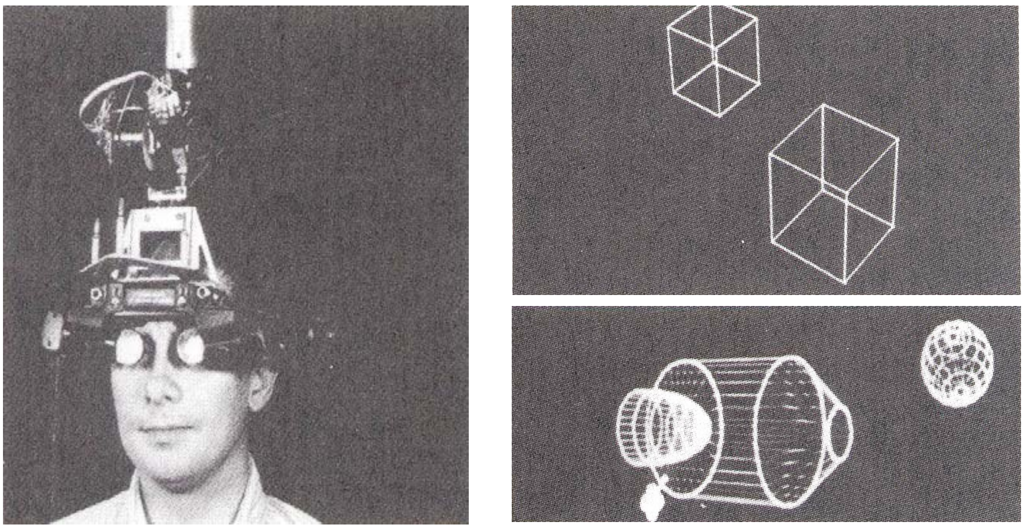


Figure 2.9: First system to display a simple synthetic image produced through computer graphics [Su65], [Su70].

Generally, one of the most significant sources of Virtual Environment technology comes from previous work in the development of realistic vehicle simulators [RS86], [Car2000], primarily for aircraft as well as for automobiles and for ships. Operators of related technology should be highly trained and since acquiring relevant training on actual settings could be either dangerous or costly, simulators could 'provide' the content and dynamics of the control environment for this purpose.

These systems are usually very expensive and the applications' properties are specialised (Figure 2.10). They have recently involved the use of Helmet Mounted Displays [Fur86], however, declining computer hardware and display costs have allowed for the implementation of personal simulators for everyday use (Figure 2.11).

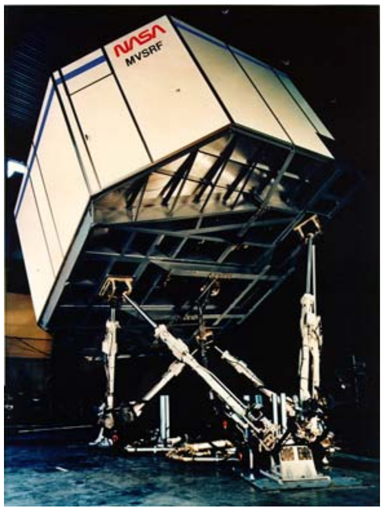


Figure 2.10: Classical Flight Simulators [RS86], [Car2000].

2.3.2 Virtual Environment Technology Issues

In recent days, according to the latest VE technology review by Frederick Brooks [Br99], Virtual Environment technologies 'barely work'. Cheaper high-resolution virtual environment systems have become commercially available and some VE applications are routinely operated for the results they produce. Following the review by Frederick Brooks [Br99], there are four technologies that are significant for VE systems:

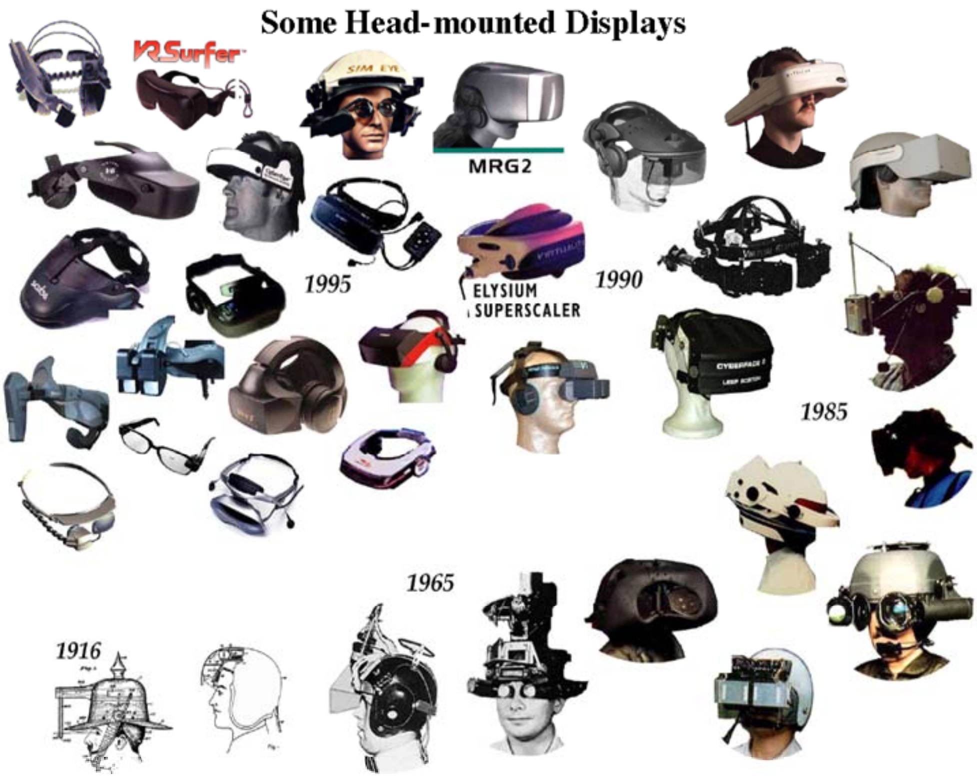


Figure 2.11: Some Head Mounted Displays [El2001].

- *The visual, aural or haptic displays* that immerse the participant in the virtual world and that block out contradictory sensory information from the real world. Display technologies include Head Mounted Displays (HMDs), CAVEs, workbenches and panoramic displays. The CAVE is an implementation of the University of Illinois (it stands for Cave Automatic Virtual Environments). It consists of a four-sided (or six-sided) ‘room’ or ‘cave’ with separate interactive computer graphics projections on each side, each one driven by one of a set of co-ordinated image generation systems [CSDKH92]. The *workbench configuration* lays a rear-projection screen flat and positions the projector so that the workbench’s length approximates that of a human body. One, two or more viewers each perceive a custom-generated image. A *panoramic display* employs one or more screens that are arranged in a panoramic configuration. This arrangement suits group viewers. Additionally, ‘fish tank virtual reality’ refers to the use of a desktop monitor to achieve interactive stereo, head-tracked scenes [AB93].

- *The graphics rendering system* that generates the interactive computer graphics imagery at a specific frame rate.
- *The tracking system* that continually reports the position and orientation of the user's head and limbs. With wide-range trackers available, though, the arrangement of related wires has become highly troublesome. There is commonly, a significant latency between user motion and its representation to the visual system.
- *The database construction* and maintenance system for building and maintaining detailed and realistic models of the virtual world.

2.4 Human Factors Issues

Virtual environments are envisioned as being systems that will increase the human-computer communication bandwidth by simulating natural human interaction through appropriate stimulation of sensory channels. This section provides an overview of human factors issues related to VE technologies [SMK98], including

- *Human performance efficiency* in the VE which could be affected by user and task characteristics, human sensory and motor physiology, multi-modal interaction, current interface and technology limitations and a need for novel interaction metaphors.
- *Health and safety issues* related to inaccurate mapping of the human perceptual system with the virtual interface of which simulator sickness is perceived as the most significant.

2.4.1 Human Performance Efficiency

Wann & Mon-Williams [WM96] stated that the goal of VE systems is to 'build (virtual) environments that minimise the learning required to operate within them, but maximise the information yield'. In order to maximise the efficiency of communication for the information conveyed in VEs, it is commonly thought that the development of specific design guidelines is required. This is not straightforward, though, as different applications and interfaces have to be accommodated.

If participants cannot perform the application tasks efficiently, the effectiveness of VE interfaces and technologies for simulation and training are compromised. The ability, for instance, of a participant to complete a set of generic tasks in a VE application may influence overall performance. The Virtual Environment Performance Battery (VEPAB) by Lampton et al. [LKGBMB94] is a move towards benchmarking VE performance. According to the authors, simple tasks are preferable to complete training scenarios since they can be employed in different applications and facilitate measurement. The VEPAB software platform includes tasks under five categories:

- Vision (acuity, colour, search, object recognition, size estimation and distance estimation)
- Locomotion (navigation, flying)
- Tracking (controlling positions of cursors)
- Object manipulation
- Reaction time tasks

The evaluation of this platform indicated that participants were sensitive to practice effects and that individual characteristics of participants should be accommodated in the design of the tasks. The results of such tests, generally, could provide a baseline of effectiveness towards the evaluation of VE implementations.

It is important to determine the *types of tasks* for which benefits can be obtained by using VE technology. The nature of the tasks being performed could directly influence how effectively participants perform them in a VE simulation. Wann & Mon-Williams [WM96] have indicated that tasks such as dimensional assessment, visual detail enlargement, design visualisation and data visualisation could acquire benefits from the VE technology. An understanding should be obtained between task characteristics and VE characteristics (for instance, stereoscopic displays, real-time interactivity, multi-sensory feedback) that effectively support specific task performance, initially within the VE and upon transfer to the real-world task. In particular, stereopsis could result in benefits in VE performance [EB94] and VE systems could present problems related to the lack of appropriate stereopsis [RMW94].

Stanney et al. [SMK98] provides a simplified categorisation of *individual differences* on VE performance, which deal with the input (interpupilary distance), throughput (cognitive or perceptual styles) or output (human performance). Characteristics that significantly influence VE experiences need to be identified in order to design VE implementations that accommodate individual needs. Significant individual differences have been noted, as mentioned above, by Lampton et al. [LKGBMB94] while evaluating their task-based benchmark software platform. In more traditional Human Computer Interaction research, one of the primary characteristics that interface designers adapt to is performance of novices vs. expert users [DFAB93]. Experience of interaction with VEs would influence the manner with which users interact with the VE and also the mental representation of a VE over time. For instance, there is an issue of how to assist participants of low spatial ability in effectively navigating VEs. Stanney & Salvendy [SS94] indicated that although participants with low spatial ability are unable to mentally represent the structure of complex systems, they are capable of doing so when the systems are well organised and when the focus is clearly defined on acquiring their structure. However, if task workload is high during the initial stages of system use, low-spatial ability participants do have difficulty generating an accurate representation of the VE layout [WWH94]. Other individual differences that could affect VE interaction could be elements of personality, age and deficits in perception and cognition. For instance, Howe & Sharkey [HS98] suggest a method for identifying those individuals who are more suited to using VE systems based on their competence (mental adaptability, spatial awareness, visual perception, co-ordination) and temperament (personality traits). This method, though, seems quite broad and not adequately related to VE. Although the authors take input from established work on personal profiling, it is not clear how the actual weights employed will have to vary for specific applications.

In certain task performance based evaluation studies it has proven quite difficult to reveal statistically significant differences between discreet technological conditions, when the only measure at hand is the performance of a task. Ruddle et al. [RPJ99] investigated large-scale VE navigation comparing a desktop display and a HMD. There was no statistical difference between the two types of display in terms of the distance that participants travelled in order to complete the task or the mean accuracy of their direction estimates. Also, Pausch et al. [PPW97] found no significant differences for visual search performance in a study comparing a head-tracked VE displayed on a HMD with a stationary HMD employing a hand-based input device.

This experimental study was repeated by Robertson et al. [RCD97] incorporating a typical desktop installation that didn't replicate the full results. Moreover, it now appears, that aftereffects of VEs can sometimes be ignored by the sufferer when he/she is confronted with tasks that must be performed. For example, astronauts experiencing space motion sickness in one study [TMPV98] showed no decrements in manual tracking or complex reaction times and performed their assigned operational tasks well. In this case, the relatively low rate of reported incidences of performance deficiencies associated with space motion sickness might result from a very high degree of astronaut selection, training and motivation, which might have compensated for motion sickness effects upon performance in this group.

VE performance measures, thus, **need to focus on more than task performance** to be effective. Due to the complex nature of simulating the human perceptual and motor mechanisms into a successful VE application, especially for simulation and training, task performance measures alone do not reflect how effective the specific implementation or hardware could be [RPJ99], [TMPV98].

2.4.2 Aftereffects

This thesis is centred on human experimentation using Head Mounted Displays (HMDs). In this section, research relevant to aftereffects associated with their usage is going to be presented. Welch stated that 'strictly speaking, any effect observed after the participant has returned to the physical world qualifies as an aftereffect' [Wel97]. While the term "cybersickness" evokes thoughts of overt nausea and general malaise during VE use, it actually comprises a number of less obvious effects. Aftereffects such as balance disturbances, visual stress, altered hand-eye co-ordination, drowsiness, fatigue, lowered arousal or mood, eyestrain, malaise and headache are a disturbing by-product of immersive VE exposure [SSetal98]. The most common VE post exposure aftereffect is disturbed locomotor and postural control. Perceptual-motor disturbances have been observed with the use of VE devices [KSOD97]. Also, proprioception (felt position) is found to be easily biased giving rise to proprioceptive illusions such that participants of VE simulations 'feel' their hands or body are in confusing locations relative to visual displacement.

Prothero [Pro98] makes a distinction between simulator sickness (the generic feeling of sickness from exposure to a computer generated world), motion sickness (components of simulator sickness which are inherent to the specific stimulus and which would be present if the simulation were a perfect representation of the real world) and interface sickness (which arises from imperfections in the technology such as lag and geometrical distortions). This is a significant categorisation since aftereffects related to ‘interface sickness’ are likely to be reduced as technology advances.

Generally, 80 to 95% of individuals exposed to a VE system report some level of post-exposure symptomatology. This is due to *participant sensitivity*, as well as *system characteristics* (e.g., stimulus strength) [SSetal98], [Nic99], [CNRW99].

Participant sensitivity

According to the Stanney & Salvendy report [SSetal98], age, gender, motion sickness histories and prior experience have been shown to be useful factors in identifying participants who are generally sensitive to provocative motion environments. For instance, those over the age of 25 are about half as susceptible to motion sickness as they were at 18 years of age. It has been found that females experience greater motion sickness than males, however, it is not currently known if these facts could be generalised to VE exposure. The age at which children should be allowed unconstrained access to Head-Mounted Display (HMD) systems is still open to debate since binocular vision develops early and is relatively stable by the early school years. Similarly, adults vary in the robustness of their visual systems. It can be predicted that someone with unstable binocular vision may experience stronger post-exposure effects if there are stimuli that place some stress on either the accommodation (focal) system, vergence system or the cross-links between them [WRM95]. To date, however, the research into visual aftereffects has been almost exclusively upon users with robust binocular vision for ethical reasons. Individual variations in inter-pupilary distance (IPD) may also have consequences for calibration of HMDs. Physical measurement and adjustments to suit individual IPDs, although necessary, are often impractical [RMW94], but such adjustments can be tuned to the user with appropriate software routines [WRM95].

Stimulus Strength

General system factors thought to influence stimulus strength include: mismatched IPDs [MWR95], possibly large field-of-view (FoV) [Art2000], update rate and lag between head movement and update of the visual display [SG95], visual simulation of action motion (i.e., vection [KBDH96], [Pro98]). Particularly, DiZio & Lackner [DL97] have identified large end-to-end visual update delays and a large FoV as significant factors that induce aftereffects when a VE is displayed on a HMD. They emphasise that experimenters must devise formal ways of measuring system characteristics [RMCH97]. According to Kennedy et al. [KSD2000] simulator sickness increases with longer exposure but also diminishes with repeated sessions. Hence, employing short, repeated exposures to manage simulator sickness may be an effective strategy.

Measuring Aftereffects

By identifying consistent measures of possible aftereffects, the development of technological or adaptive countermeasures would be facilitated. As analysed, the user sensitivity rate related to VE aftereffects tends to be related to user susceptibility, as well as system characteristics (e.g., stimulus strength). Thus, in order to determine the differential effect of user and system factors on symptomatology it is important to determine the various drivers of cybersickness and to create a predictive model whereby cybersickness and aftereffects can be foreseen.

Subjective measures of aftereffects mainly include self-report of symtomatology in the form of questionnaires administered after the exposure to the VE application. The most commonly used tool is the Simulator Sickness Questionnaire (SSQ) [KLBL93] which was originally devised to evaluate aircraft simulator systems. The SSQ consists of a checklist of 16 symptoms, each of which is designated in terms of degree of severity (none, slight, moderate, severe). A weighted scoring procedure is used to obtain a global index, known as the Total Severity (TS) score which reflects the overall total discomfort level. The SSQ also provides scores on three subscales representing separable dimensions of simulator sickness (i.e., nausea [N], oculomotor disturbance [O], and disorientation [D]). By means of these subscales, cybersickness can be more systematically characterised. Participants report the degree to which they experience each of the above symptoms as one of ‘none’, ‘slight’, ‘moderate’ and ‘severe’. These are scored respectively as 0,1,2,3. To compute the scale scores for each group, the reported value for each symptom is multiplied by the weight each symptom is assigned and then summed. The total SSQ score is obtained by adding the scale scores across the total score for each of the three components and multiplied by 3.74. Weighted scale scores for each column, individually, can be calculated by multiplying the ‘nausea’ scale score by 9.54; the ‘oculomotor’ total score by 7.58; the ‘disorientation’ total score by 9.54. For a good review of other related subjective measures see [SSetal98].

Using the SSQ as the measure, Figure 2.12 shows the average Total Severity (TS) scores for cybersickness (for aftereffects related to VE/HMD exposure), simulator sickness (for aftereffects related to simulators exposure), and space sickness (for aftereffects related to space exposure) [KS97], [SSetal98].

An examination of immediate post-exposure profiles of VE systems using HMDs indicate that, on average, these systems tend to produce more disorientation [D] than nausea [N] symptoms and least of oculomotor-related [O] disturbances. This means that users of these systems are more prone to experience dizziness, vertigo, general discomfort, increased salivation, sweating, and nausea, than they are to encounter headaches, eyestrain, or difficulty in focusing. This D>N>O profile does not match the profiles of other provocative environments, including simulator sickness which consists of an O>N>D profile, and space sickness which consists of an N>D>O profile (Figure 2.13). Since flight simulators and VEs are both visually interactive environments one might expect their symptom profiles to match. Their diverse profiles indicate quite convincingly, however, that these systems differ substantially in the symptoms they produce. This may indicate that a new factor analysis is required to optimise the use of the SSQ for VE systems. Factor analysis could identify new clusters for symptoms.

Other methods for measuring problematic adaptation to VE systems could evolve from postural instability or unsteadiness post exposure tests. Postural equilibrium has been proposed as an indexing system for identification of VE systems that could cause severe aftereffects [DL97].

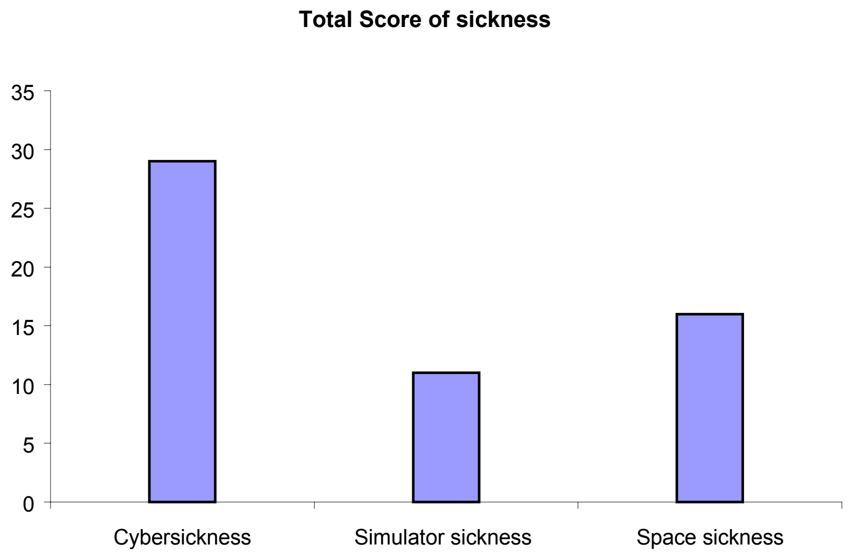


Figure 2.12: Average SSQ TS score for Cybersickness, Simulator sickness and Space sickness [KS97].

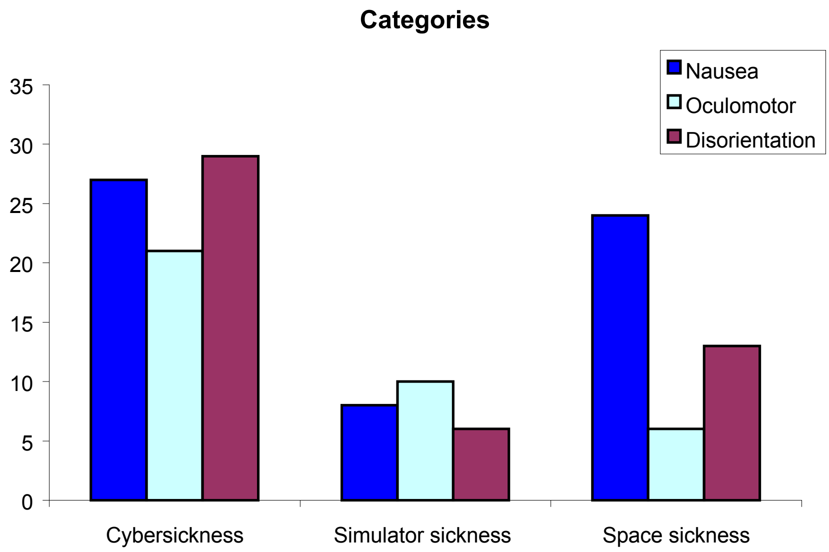


Figure 2.13: Average SSQ scores of Nausea, Oculomotor and Disorientation subscales for Cybesickness, Simulator sickness and Space sickness [KLLBH92].

DiZio & Lackner [DL97] have measured postural instability after VE exposure, using a Kistler force platform and the standard Romberg posture. They found a five-fold increase in sway amplitude after only 15 minutes of VE exposure as compared to pre-exposure levels. Partisipants' level of aftereffects were measured without the HMD on 5-minute intervals post exposure. These aftereffects dissipated exponentially to baseline and were no longer measurable 15 minutes after VE exposure. Cobb & Nichols [CN98] revealed a positive correlation between self-reported simulator sickness symptoms and balance-related symptoms but no correlation between the former and performance measures to postural instability. Relevant work has been conducted by Prothero [Pro98], in relation to vection (visually induced illusion of self-motion). He introduces the Rest Frame Construct (RFC) based on the premise that spatial judgements are made with respect to a 'rest frame' that is not physically determined but carefully maintained by the nervous system and could be related to the reduction of simulator sickness. Additionally, measuring changes in hand-eye co-ordination, visual functioning, heart rate levels are perceived significant towards predicting post-exposure aftereffects.

The relationship between aftereffects of VEs and task performance is not fully understood or generally concluded [SSetal98].

2.5 Simulation Fidelity for Dynamic Scenes

Simulation fidelity is characterised as the extent to which a Virtual Environment (VE) and interactions with it are indistinguishable from a real environment and relevant interactions [WHK98]. The mapping from the real world environment to the computer graphics environment is mediated by *environmenta*l or *visual* fidelity. The term *visual fidelity* refers to the degree to which visual features in the virtual environment conform to visual features in the real environment. *Interface or interaction fidelity* refers to the degree to which the simulator technology (visual and motor) is perceived by the trainee to duplicate the operational equipment and the actual task situation. Training, for instance, in a VE with maximum fidelity would yield transfer equivalent to real world training since the two environments would be indistinguishable. Increases in fidelity, though, could prove to be computationally demanding. The effectiveness of VE technology results in a trade-off between economic and technological variables, related to equipment and computational power. It is, therefore, important to gain an understanding of which technological or computational variables could be sacrificed without degrading performance.

This thesis is focused on simulation fidelity metrics that are based on a methodology (presented in Chapter 3) which takes into account the cognitive processes participants employ to achieve a task in a real world task situation in comparison to its 3D simulation counterpart. In this section, research related to the sense of ‘presence’ is reviewed since presence is perceived as a metric for VE effectiveness and fidelity that is not linked to task performance. Importantly, in section 2.5.4, research that compares a real world task situation to its 3D simulation counterpart is going to be presented in detail.

2.5.1 Usability Engineering for Virtual Environment Applications

Usability has gone far from being an issue only in the design of PC-based software applications and office-based peripherals and devices to a central issue in the design of a vast range of technologies, particularly handheld and mobile personal systems and virtual and augmented reality applications. There is now a question mark whether established usability engineering methodologies will be suited to emerging technologies. Various studies have been conducted to assess the simulation fidelity of computer graphics imagery and relevant interface technology employed for a specific VE application [GHS99], [Gab97]. These studies have been almost uniformly focused on human performance efficiency in the VE, since for an application to be successful, this should be maximised and often compared to the real world. Gabbard et al. [GHS99], [GSRH99] presented a methodology of evaluating a virtual battlefield VE based on

- User task analysis: identifying a complete description of tasks and subtasks required to use a system.
- Expert guidelines based evaluation: identifying potential usability problems by comparing a user interaction design to established usability design guidelines.
- Formative user-centred evaluation: observational evaluation method that includes participants testing the design from early stages.

- Summative comparative evaluation: empirical assessment of an interaction design in comparison with other maturing interaction designs for performing the same user tasks.

This strategy borrows elements of traditional evaluation techniques for Graphical User Interfaces (GUIs) evolved for the evaluation of VE systems. Generally, in order to maximise the efficiency of a VE implementation it is often stated that a set of 'design principles' should be identified which would enable intuitive interaction with the VE. I doubt if this, often content-free requirement is ever possible or useful since application goals differ and generic guidelines normally do not address specific issues. Due to the complex nature of human-VE interaction, VE performance measures need to focus on more than task outcome to be effective.

Generally, usability has started to emerge as a necessary component of specific technology implementations. In recent literature it appears to be closely interconnected to the design of the system or interaction method in question [SSK2001], [LST2001], [BW2001], [MEBS2001].

2.5.2 The Notion of 'Presence' as a Method of Evaluation for VEs

A taxonomy of VEs is based on three main components: *autonomy*, e.g. computational models and ability of processes to act and react to simulated events, *interaction*, e.g. software architecture of the human-machine interface of the virtual environment system and *presence* [Zel92] (Figure 2.14). What sets VE technology apart from its ancestors is that in VE systems users can receive a number of distinct multi-sensory stimuli (i.e., visual, auditory, haptic) which are intended to provide a sensation of 'natural' interaction with the virtual environment, and, consequently, an illusion of being 'present' in a VE. However, a distinction between the notion of *presence* in a VE and *technological immersion* needs to be drawn as, for example, distinguished by Slater & Wilbur [SW95].

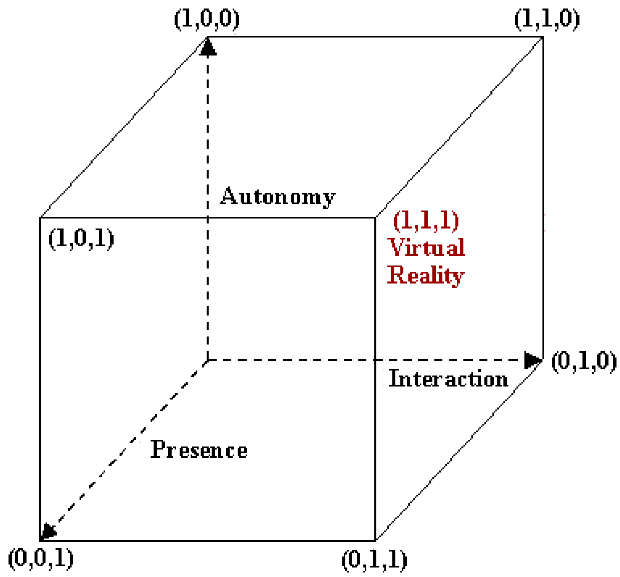


Figure 2.14: The AIP cube [Zel92].

According to this research [SW95] immersion is perceived as a quantifiable description of technology, mainly determined by the extent to which visual displays are (a) inclusive (b) extensive (c) surrounding and (d) vivid, as opposed to presence which is a 'state of consciousness'. From a different point of view, Singer & Witmer describe immersion as an individual's perception and reaction to a VE [SW99] referred here as *psychological immersion*. Pausch et al. [PPW97] relates presence directly with technology attributes such as the 'immersiveness' of a HMD by quantifying immersion in terms of the performance of a task. This is quite a simplified approach since, generally, an obvious or straightforward relationship between presence and immersion has not been proven. Technological immersion could influence presence and task performance in a VE in unknown ways. The notion of presence is perceived as a construct that will provide a tool for the evaluation of VE software platforms, interaction metaphors and applications. Presence does not depend directly on task performance and can therefore be applied to different task situations. Its definition and measurement is still a challenge for the VE field and evokes fascinating discussions and research [SSetal98]. Its usefulness, though, remains to be proven.

2.5.2.1 Definitions

'Presence' generally, refers to the sense of being present in time or space in a particular location (Webster's II Dictionary, 1984). In the world of media and emergent technologies such as video conferencing, high definition television and home theatre, presence is defined as the *perceptual illusion of non-mediation* [LD97]. An 'illusion of non-mediation' occurs when the user fails to perceive the existence of a medium in his/her communication environment and reacts as he/she would if the medium were not there [LD97]. Presence, in this sense, can occur in two particular ways: (a) the medium can appear transparent and function as a large open window with the medium participant and the medium content in the same physical environment (b) the medium can appear to be transformed into something other than a medium, into a social entity. This definition of presence, it is argued, can be applied to any medium and encompasses the following six conceptualisations:

- Presence as social richness: The medium is perceived as sociable, warm, sensitive when it is used to interact with other people.
- Presence as realism: The degree to which a medium can produce accurate representations of objects, events and people.
- Presence as transportation: The participant is 'transported' to another place or other places are transported to the participant or more than one participants are transported together to a place that they share.
- Presence as psychological and technological immersion: When a medium submerges the perceptual system of the participant.
- Presence as social actor within medium: Medium personalities as TV personalities use direct access via the medium to generate a sense of real-life interaction.
- Presence as medium as social actor: This involves social responses of media participants towards the medium itself that resemble real-life interaction.

The above conceptualisation has created a body of work related to presence and media technologies. For a very good review, see [LD97] and also the work of Freeman et al. [FAPI99].

Research into VEs addresses the definition of presence from a different point of view. As VEs, in every form, potentially provide a new communication medium for human-machine interaction [El94], presence can be explained as the participant's sense of 'being there' in a VE. This is the degree to which the users feel that they are somewhere other than they physically are while experiencing a computer generated simulation [BZSS95], [HD92], [WS98], [SSMM98], [BH95]. Zahoric & Jenison [ZJ98] give a definition of presence as a '…tantamount to successfully supported action in the environment' linking the notion with task performance in the VE and less with issues of photorealistic visuals and appearances. Telepresence is defined by Sheridan [She92] and Draper et al. [DKU98] as the ability of a human operator of for instance, a robotic arm to receive sufficient information about the task environment that the human operator feels present at the remote site. Heeter [Hee92] perceives presence as a virtual experience that convinces participants that they are 'there'. This definition does not focus on how close a virtual world mimics real world sensations.

Examples of factors which research implies increase the sense of presence include many which one could, intuitively argue make an interface more 'natural'. Some examples are interactivity and pictorial realism [WBLMS96], update rate [BH95], [WBLMS96], large Field-of-View [Art2000], meaning related to the task [Pro98], stereo imagery and head tracking [HB96a] and spatialised sound [HB96b].

Generally, the degree of 'reality' necessary for a 'representation' of any kind is not a new research question. Aristotle [Ari] pointed out in his theory of unity that only details that are necessary to communicate a 'story' should be included. He also argues that a representation (referring to poetry) should not be judged according to rational rules but with perceptual ones; something real is not what objectively is true but what offers the impression of reality by convincing and emotionally engaging as being true (initial quote of this thesis). Also, the theory of theatre tells of an ongoing struggle between different representations of fictional worlds. These methods strongly affect the type of 'engagement' an audience experience. For instance, Ibsen was developing theatre to emulate reality as close as possible; this is known as naturalism. Brecht, on the other hand, developed what he called 'alienation technique'; he did not believe that the audience should be brought to believe that plays were a sort of reality, but that theatre should engage the audience intellectually. Tallyn [Tal2001] in her Ph.D thesis states in relation to that:

> 'In the end, the type of representation chosen depends on the type of engagement you want to create, i.e. if very intellectual engagement in real world issues is required perhaps it is better not to distract with fantastic representations. The creation of a fictional space is more about creating a successful closed system that makes sense within itself, rather than creating a fantasy that transports us away from reality'.

2.5.2.2 Measuring Presence

Different ways of measuring presence have been employed by various researchers. One of the 'hot, open challenges' is to measure 'the degree of presence and its operational effectiveness' as quoted in [Br99]. The development of metrics for presence could provide a conceptual and analytical framework through which to identify areas for future VE research. Generally, Ellis [El96] has obtained empirical evidence to suggest that subjective responses to the characteristics of VE systems may meet proposed criteria for explanatory constructs including repeatability, reliability, and robustness.

Subjective Measures

Freeman et al. [FAPI99] employed a form of direct subjective evaluation where users were required to provide a continuous rating of their sense of presence using a hand-held slider. Slater et al. have recently introduced a measure of presence based on self-report of 'Breaks in Presence' (BIPs) while a participant experienced a VE simulation [SS2000]. Snow et al. [SW98] has introduced a method of magnitude estimation where the participants rate their level of presence by quoting a single number. Every subsequent report of an estimate relevant to their level of presence is associated to this initial estimation.

Another way of measuring presence, in an effort to introduce a quantitative strategy, was proposed by Schloerb [Sch95]. This method is based on a participant's inability to discriminate between a real and a VE and proposed the addition of certain types of 'noise' to a real image until it is impossible to be distinguished from the virtual image. Since, with present VE technology, it is very unlikely that observers will mistake a given VE for a real one, a method for quantifying presence produced by the VE is to 'degrade' perception of the real scene until the observer can no longer differentiate it from the virtual one. The amount of degradation required to reach this level serves as a measure of the degree of presence. The primary advantage of this 'paired comparisons' method is that it does not require investigators to explain to their participants what they mean by presence or even to introduce the term at all. This is desirable because the act of carefully defining this concept for participants may create a situation in which the investigators guarantee confirmation of their hypothesis, particularly if their definition mentions or even implies the variables that they will be manipulating. This method is often perceived as a 'Turing Test' for VEs [Shl95].

The most common way of measuring presence is post-experiment self-report through questionnaires. Witmer & Singer, presented a Presence Questionnaire (PQ) and an Immersive Tendencies Questionnaire (ITQ) [WS98]. According to the authors, the PQ measures the degree to which individuals experience presence by taking into account possible contributing factors such as control, sensory, distraction and realism. The ITQ questionnaire tests the capability of individuals to be involved, engaged or immersed in every day situations that are not necessarily related to technology, investigating individual characteristics of the users. Also, Barfield & Hendrix [BH95] have adopted questionnaires that include questions in two categories: Presence and Fidelity of interaction.

Slater et al. [SSMM98] introduced also a presence questionnaire. The participant rates the extent during the experience that the particular 'space' of the application is the dominant reality as well as their level of perceiving the VE as a 'locality' or a 'place' that was visited rather than merely seen. This questionnaire has been employed in the experimental work relevant to this thesis to assess participants' level of presence and it relates to the definition of the notion itself. It also consists of a smaller number of questions and, thus, is easier to administer compared to the Witmer & Singer [WS98] questionnaire. It includes the following questions taken from [SSMM98] (these questions slightly vary to reflect the context of a specific set of experiments):

- Please rate your sense of being in the seminar room, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place.

- To what extent were there times during the experience when the seminar room was the reality for you?
- When you think back about your experience, do you think of the seminar room more as images that you saw, or more as somewhere that you visited?
- During the time of the experience, which was the strongest on the whole, your sense of being in the seminar room or of being elsewhere?
- Consider your memory of attending the seminar. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By ‘structure of the memory’ consider things like the extent to which you have a visual memory of the field, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic in your imagination, and other such structural elements.
- During the time of the experience, did you often think to yourself that you were actually in the seminar room?

Questionnaires are not an ideal research method. They can be troublesome to validate but also they are assessing the level of perceived presence after it has occurred since the participant is requested to think back to his/her experience. While subjective ratings scales are effective means of assessing presence, it is important to note that such scales should be used judiciously due to inconsistencies across different raters or rating situations [El96]. Recent research debates the structure of presence questionnaires as well exploring the defining elements of the notion itself [Sla99], [SW99].

Objective measures

Loomis [Loo92] and Sheridan [She92] observed human response to events that in the natural world would provoke ‘reflex’ reactions. If one is sitting in front of a screen and experiences a scene of a car moving towards him/her very fast, then he/she might be ‘forced’ to turn to the right or left, in order to avoid ‘collision’, responding to the moving image as if it was occurring in reality. Measures of presence could include both neurophysiological responses and reflexive motor acts. Physiometric measures to events in a VE could include posture, muscular tension, cardiovascular behaviour, and ocular responses [BW93].

Objective measures will be most useful when they are tailored to the experiences participants are intended to have in a VE system. For example, Lackner & DiZio [LD94] have demonstrated a useful automatic motor response as a measure of presence in a VE that could induce a sense of body rotation. Participants who are physically rotating (in a rotating room) but who feel stationary make errors when pointing to targets. The paths and endpoints of their movements are deviated in the direction of the transient inertial Coriolis forces generated by their arm movements. By contrast, participants who reach while voluntarily turning their torso reach accurately even though their arm movements generate large Coriolis forces. This pattern suggests that the nervous system, in generating motor plans for reaching movements, takes into account whether the body is rotating and automatically includes forces counter to the expected Coriolis forces.

Also, physiological measures such as electrodermal activity and skin temperature have been employed [Mee2000]. Findings show that such measures positively correlate with traditional presence questionnaire responses. In the study by Meehan [Mee2000] skin conductance data correlated positively with the Slater et al. [SSMM98] presence questionnaire dataset.

2.5.2.3 Relationship between Presence and Task Performance/Motion Sickness

Presence and Task Performance

It is often argued that a possible correlation of presence with task performance could be valuable towards the design of simulation and training applications. The degree to which presence can provide a means to assess the overall effectiveness of an interface or a method of interaction depends on the direct link of presence and effectiveness and could maybe established empirically. The potential of VEs for training of real world tasks is generally attributed to the level of presence provided by the VE system [She92], [Zel92]. One of the strongest challenges for presence related research is clarifying the relationship between presence and task performance [BZSS95]. Alternatively, though, the sense of presence may not be necessary towards achieving effective human performance in VEs and may even merely constitute an epiphenomenon of a VE experience [El96]. Slater et al. [SULK96] argue that:

> 'It is posing the wrong question to consider whether presence per se facilitates task performance. Rather presence brings into play 'natural' reactions to a situation (which may or may not have something to do with efficiency of task performance) and the greater the extent to which these natural reactions can be brought into play the greater that presence is facilitated'.

While it is commonly thought that performance in a high-level-of-presence VE is likely to be better than in a low-level-of-presence VE, there is little systematic research available to substantiate such a claim [SSetal98]. Witmer & Singer reported a positive relationship in one study [WS94]. This was not consistently reliable in another study by Bailey & Witmer [BW94]. Snow [Sno96] explored this relationship in his Ph.D. thesis by manipulating independent variables such as update rate, resolution, FoV, sound, textures, head tracking with participants performing a set of five tasks. Perceived presence was assessed using the technique of free-modulus magnitude estimation. A positive relationship was found between perceived presence and task performance but this relationship was relatively weak. Prior to establishing a complete understanding of the interrelationships between presence and task performance, it is not wise to state that designers of VE systems should directly pursue implementations that engender a high level of presence. Even if it is clearly demonstrated that presence influences performance, its effect may not always be beneficial. Ellis [El96] provided two examples in which analysis of the desired causal control between system and user determined that the optimal sensorimotor transformations required by the task would be achieved by placing the user outside of the VE (exocentric view) rather than the more 'natural' position of being immersed within (egocentric perspective) the VE. Even if there *is* a causal association between presence and performance (I do not believe there is), it may be a negative one.

The commonly thought positive between presence and task performance relationship (not supported though by literature as shown above) is likely to be highly task dependent. This relationship, as mentioned, is often considered to be causal, but the only evidence that one can find to support this is that it is correlational [Wel99]. This means that the variables that increase presence might also increase task performance independently of their effect on presence. Therefore, an unambiguous test of whether presence is causally related to performance requires that presence be manipulated in a manner that is unlikely to directly influence performance. For example, Welch [Wel97], [Wel99] suggested conducting a study in which the sense of presence is manipulated by means of an element of the application that is not directly related to the task. For instance, this could be ambient sound while using a visual-motor task to measure performance. Thus, since the variation in sound levels is unlikely to have a direct effect on visual-motor behaviour, but is assumed to increase the sense of presence, an effect of this manipulation on performance is likely to be the causal result of the increased presence.

Presence and Motion Sickness

Possible associations between presence and aftereffects may aid investigation into how to more effectively design the human-VE interface which may help eliminating such maladies. It has been hypothesised that the more accurate a VE recreates the physical world and the observer's sensory-motor relationship to it, the stronger the sense of presence and the less the aftereffects [Wel97]. Ideally, when the VE involves an accurate visual and motor representation of the real world and relevant interactions with it, there will be no adaptation process and therefore no post-exposure aftereffects and also, a strong sense of presence. If, as is more likely, a VE creates an imperfect version of real world experiences and initiates an adaptive process to compensate for these imperfections, users will have little or no initial problem discriminating it from the real world (i.e., will experience less than maximal presence) and will reveal a post-exposure aftereffect as a result of this adaptation [SSetal98]. The strength of the initial sense of presence in a VE will be negatively correlated with the aftereffects produced by it. There is some empirical evidence to support this relationship. Singer et al. [SAMG97] reported a negative relationship between sickness (as measured by the SSQ [KLBL93]) and presence (as measured by the PQ [WS98]), although this relationship was not always significant. There may be evidence that contradict these results as well. Wilson et al. [WNH97] reported a positive relationship between sickness (as measured by a Short Symptom Checklist) and presence (as assessed by a subjective questionnaire, secondary task and observational measures). One should, however, note the different measurement tools that were used, a fact that may have contributed to the conflicting nature of the results.

Substantial adaptation to a VE may produce a strong sense of presence, which suggests a second prediction [Wel97]. The greater the aftereffects from a VE, the greater the increase in presence over the period of exposure to that VE. This hypothesis is based on the assumption that sizeable VE aftereffects are the result of sizeable VE adaptation. With this adaptation, observers will perceive much less of the intersensory application shortcomings that initiated the adaptive process early on and which, at that time, contributed against a strong sense of presence.

Finally, there is the possibility that aftereffects and presence are both correlated (either positively or negatively) with a third factor or variable, for instance, vection. Changes in one could initiate changes in the other but solely via the intervening factor, and thus, there is no direct connection between them [Pro98]. This is the equivalent of the Welch [Wel97] argument for the relationship between presence and task performance.

2.5.3 Subjective Responses to Lighting

Light has the obvious function of providing visibility for visual task performance. Flynn [Fly77] argues that lighting properties should begin with the overall user well being, the visual quality of a room and should not be limited to task visibility. The author and his associates had been investigating evidence that views human responses to spatial lighting or visual patterns as, to some extent, shared experiences. The importance of vision is, in part, an experience of recognising and assimilating communicative patterns. For instance, white lanes provide a spatial limit in roads. James Gibson has explored the idea of spatial meaning and information content [Gib71]. He has suggested that

> 'the optic array from a picture and the optic array from the real world can provide the same information without providing the same stimulation. Hence, an artist can capture the information about something without replicating its sensations'.

He argues about a new theory of visual perception based on the idea that light can convey information; subsequently, the brain constructs the phenomenal mental image depending on this information.

Flynn [Fly77] noted that many lighting systems are designed merely to function in a 'permissive' way, i.e. simply to permit performance or participation in some activity that involves vision, without attempting to influence participants impressions or behaviour. Many lighting designs, however, especially in a commercial context, may intentionally or unintentionally function more actively as selective intervention in human visual experiences: focusing attention, guiding circulation and otherwise affecting impressions of a room or activity. The author suggests a procedure for investigating the effect of light on impressions and behaviour based on the use of Semantic Differential (SD) rating scales, including adjectives such as 'clear-hazy', 'pleasant-unpleasant', etc. Work with such scales has identified several broad categories of impression that can be applied to lighting. The categories of impression of particular interest are:

- Perceptual categories such as visual clarity, spaciousness, spatial complexity, colour tone, glare.
- Behaviour setting categories such as public vs. private space, impressions of relaxing vs. tense space.
- Overall preference impressions such as impressions of like vs. dislike or impressions of pleasantness.

It is interesting that investigations of similar light settings in different rooms and with different *object arrangements* or *activity settings* indicates that the modifying effect of lighting is consistent across rooms [Fly75]. This reinforces the theory that these subjective impressions are more a function of the actual lighting characteristics than the actual environment in question.

A second element of the procedure to assess subjective responses to lighting involves the use of Multi-Dimensional Scaling (MDS). These dimensions or modes of lighting are:

- The Overhead Peripheral Mode, referring to a lighting of vertical surfaces as distinguished from overhead lights that illuminate central horizontal surfaces.
- The Uniform/Non-Uniform Mode, referring to the appearance of the room or of major surfaces in the room as well as referring to the appearances of objects and artefacts within the room.
- The Bright/Dim Mode, referring to the perceived intensity of light.

- The visually Warm/Cool Mode, referring to the perceived colour tone of the light in the room.

Generally, the subjective responses to lighting procedures indicate an effort towards assessing lighting designs from an impression point of view rather than a task point of view. One could argue that the presence related research for VE technologies is striving to achieve similar goals: assess a software platform or a virtual interface generically, not by necessary linking this assessment with task performance although the distinct relationship between presence and task performance is often considered crucial. It would be interesting to apply the subjective impressions lighting procedure to assess the simulation fidelity of a computer graphics scene vs. its real counterpart. The final goal of a significant paper by Rushmeier et al. [RGPSR95] on perceptual image quality metrics that was not completed[1] was to relate subjective impressions of an environment to values computed from measured luminance images. Additionally, it would be interesting to identify if there is any correlation of this particular set of data[1] with perceived presence (investigated in Chapter 5).

I was introduced to a new concept called Kansei Engineering in a private discussion[2] with the following example paraphrased: When we enter a building, we mentally create an overall impression. We could in this process make statements such as 'I really like this building but I am not sure why, I can't pin it down'. This pleasant feeling of liking or disliking the 'aura' of a space, since it is totally subjective, can't be formulated easily into a concrete framework. It exists, though, and a new field named Kansei Engineering is concerned with its investigation. Kansei Engineering is a technique for product development [KMN94] which takes into account the desirable features of products as perceived by end users themselves. Kansei is a Japanese term that means psychological feeling or image of a product. It refers to the translation of consumers' psychological feeling about a product into perceptual design elements. Kansei engineering is also sometimes referred to a 'sensory engineering' or 'emotional usability.' This technique involves determining which sensory attributes elicit particular subjective responses from people and then designing a product using the attributes that elicit the desired responses. The technique involves building a database of the keywords that represent consumers' feelings towards products and

[1] Personal communication with Dr. Holly Rushmeier, 2000.
[2] Personal communication with Prof. Thomas A. Furness III, 2000.

then these are used to produce scales, which are then employed to evaluate a number of products. These subjective responses can be assessed using sets of bipolar attribute rating scales. A typical bipolar attribute rating scale uses a pair of opposed terms, such as simple vs. complex or enticing vs. repulsive, placed on a continuum represented as a line. Factor analysis techniques are then used to identify those features of product designs that correlate with consumers' feelings. The technique is also being developed to produce design rules for the development of products, and in the evaluation of prototype design solutions. It is reported that the technique has been applied in Japan to the design of cars, construction vehicles, costumes, houses, etc., and has also been used in Korea, China and the USA.

The transfer of participant overall feelings and impressions about light and spaces relevant to the theories of subjective responses to lighting as well as Kansei Engineering to the computer graphics rendering of a space, is a challenge that could only be fulfilled if an understanding of the human perceptual systems is achieved. Still, this is not a straightforward issue by any means since it is not clear how impressions as such influence performance in the real or in a computer graphics world. A questionnaire that reflects subjective responses to lighting as analysed in this section is going to be incorporated in the experimental design employed in the Chapter 5 of this thesis.

2.5.4 Comparing Real and Simulated Virtual Environments

Generally, VE applications could fall into two categories: Realistic Virtual Worlds and Magical Virtual Worlds [Smi87]. Realistic Virtual Worlds are meant as simulations of the real world, mainly for training purposes such as, for instance, flight simulators. Magical Virtual Worlds support activities that don't follow 'reality' rules and have no implications outside of the virtual domain, including, for example, data visualisation. Although, this is quite a broad and simplified categorisation with often shady boundaries, it is a significant separation to be considered for VE usability engineering research. Realistic VE must take account of means over ends towards achieving the application goals for simulations that are meant for transfer of training in the real world. However, Magical VR is less constrained by training effects and can therefore supply techniques which are impossible in the real world, for instance, exploiting 'lack of gravity' in a VE or pointing and teleporting to remote objects.

For Realistic VEs, one way of getting an objective baseline for effectiveness of an application is to evaluate that against the real world. Various VE techniques could be compared against that baseline. One could argue that, generally, matching real world performance is not the ultimate goal of VE research as it runs the risk of limiting the implementation of imaginative interaction techniques and design in order to match the limitations of the real world. However, when a VE application requires a high level of simulation fidelity, comparing against the real world in a controlled way, proves to be a useful benchmark. This methodology is the an important element of this thesis. A collection of studies that compare an interactive computer graphics scene to its real world counterpart will be presented in this section in detail.

Nemire et al [NJE94] argued that measurements within a VE application could be compared with those in the corresponding real environment to determine how effectively the VE communicates the conditions of the real environment. This study investigated spatial orientation in a virtual environment and compared results with those obtained from an analogous experiment performed in a physical environment. The authors argue that this comparison provides a measure for assessing how a virtual environment promotes the spatially related physiological and psychological responses ordinarily obtained in a physical environment. In particular, the influence of a pitched optic array on the perception of gravity referenced eye level (GREL) was investigated. Different spatial references may lead to different judgements of eye level. Results indicated that a physical array biased GREL more than the equivalent geometrically identical pitched array. The addition of two sets of orthogonal parallel lines to the virtual pitched array resulted in the same bias as that obtained with the physical pitched array. They concluded that knowing which display parameters are essential for task performance will enable software designers to make informed decisions about the level of abstraction that is necessary for performance, thereby reserving computational and display resources for information that impacts the most knowledge about the scene. The participants in this study had a slaved head position.

Mizell et al. [MJSP2000] investigated whether immersive VE enhance performance over more conventional displays when visualising complex geometry for transfer of training in the real world. Participants of a first experimental study were shown an abstract rod sculpture in real-life as well as on a typical display with a joystick and on a head tracked HMD. The participants had to physically demonstrate their understanding of a complex structure by trying to make a real-life duplicate based on viewing the real-life model or a 3D model under the different conditions. Results showed a significant superiority of the physical representation over the two computer graphics representations, but no statistically significant difference between the monitor-joystick and the HMD head tracked conditions. Because of certain confounds on the experimental design, a second study was conducted. The viewing conditions were monocular or stereo imagery, different types of tracking and varied size of the geometry displayed, all conducted in a CAVE. The head-tracked immersive VE was shown to provide a statistically significant advantage over joystick interfaces, especially in the case where the displayed structure was shown in super scale surrounding the subject. There was no difference between stereo and mono imagery. This indicates that technological immersion makes a difference when the geometry is in some sense 'immersive' since the natural way to visualise a set of geometry is to be surrounded by it.

A study conducted by Slater et al. [SSUS2000] compares behaviour in small groups. The group members perform a task in a VE and then continue the same task in a similar real world environment. The purpose of the experiment was to compare various aspects of the social interactions among the group members in the two environments and didn't examine task performance. Ten groups of three people who had never met before were introduced in a shared VE and carried out a task that required locating puzzles on the walls of the 3D space and subsequently solving them. The group then continued the same task in the real world. In each group, there was one participant wearing a HMD with head tracking and two participants who experienced the same environment displayed on a typical monitor. The results suggest that the participant wearing the HMD tended to emerge as the leader of the group in the VE, but not in the real meeting that followed after the VE exposure. The study also revealed a positive correlation between presence and co-presence (the sense of 'being' *with* other people).

Witmer & Sadowski [WS98] also suggested that comparing human performance in a VE with performance in the real world could indicate which aspects of VE technology need improvement. Using a technique validated to measure real world distance judgements accurately, the authors compared relevant performance in a real world environment with performance employing a 3D model of that environment. The participants in the VE condition were placed on a treadmill using a stereoscopic BOOM-type monochromic head-tracked display. This technique required participants to walk without vision to a target after viewing it for 10 seconds. VE accurate distance judgements averaged 85% of the target distance whereas real world judgements averaged 92%. The magnitude of the relative errors in the VEs was twice that in the real world indicating that the VE degraded distance judgements. The VE was rendered low quality monochrome and, thus, differed substantially to the illumination in the real world.

A study was carried out by Usoh et al. [UCAS2000] to assess if two questionnaires, widely employed by the research community to assess the level of presence [SSMM98], [WS98], could differentiate between real and virtual experiences. It is argued that relevant questionnaires should score higher for the real experiences compared to their computer graphics counterparts. One group of 10 participants searched for a box in a real environment. Another group of 10 participants used a 3D replica of the environment to carry out the same task. Both of the questionnaires did not reveal an overall statistically significant effect across conditions. The Slater et al. presence questionnaire [SSMM98] showed a significant difference in two out of the six questions related to presence in the questionnaire. A similar methodology of comparing a real world application with its computer graphics simulation counterpart is adopted here towards validating a measure of presence.

The purpose of the investigation by Witmer et al. [WBK96] was to evaluate how well a 3D replica of a complex office building displayed on a head-tracked HMD trained participants to navigate in the actual real world building simulated. Sixty participants studied route directions and landmark photographs, then rehearsed the route using either the 3D model, the actual building or verbal directions and photographs. The participants' route knowledge was then assessed in the actual building. The results indicated that VE training produced superior route knowledge than verbal training, but less than the training in the actual building. It was argued that VEs adequately representing real world complexity can be employed as training media towards learning complex routes in buildings and should be considered when the real world space is not suited for training.

Relevant studies focused on spatial perception will be thoroughly analysed in Chapter 3. As shown in this section, the research community has resulted in comparing real world task situations to computer graphics simulations for reasons ranging from comparisons of performance towards evaluating a system, validation of presence measures and qualitative assessment of participants' behaviour in shared VEs experiences. This is proven to be a useful process.

2.6 Summary

This chapter introduced a set of fundamental terms in computer graphics starting with defining light and its properties, light energy, photometry and radiometry. Subsequently, computer graphics illumination models were analysed. The following sections were focused on visual perception and its application to computer graphics rendering. Perceptually-driven rendering mechanisms as well as perceptually-based image quality metrics were presented. These employ knowledge of aspects of the Human Visual System (HVS) to achieve faster, more efficient as well as higher quality rendering algorithms in an effort to produce synthetic images of a perceptual equivalence to their real scenes counterparts. Virtual Environment technologies are built on top of this underlying mechanism, introducing a complex interface which includes motor as well as visual simulation of human interaction. To conclude relevant research on image synthesis and various psychophysical studies comparing a computer graphics rendering to its real scene counterpart were presented. This set of studies involves *static*, 2D scenes.

An overview of human factors issues for VE technologies such as Head Mounted Displays (HMDs) follows, focused on human performance efficiency and health/safety issues presenting ways of measuring the level of their occurrence. This analysis concludes with a reference to usability engineering methods for VE applications.

The notion of ‘presence’, defined as the ‘sense of being there’, is perceived as a construct which will aid the evaluation of specific designs employed for VE applications. Metrics for presence include subjective and objective methods. The relationship of perceived presence with task performance and aftereffects is not yet determined and relevant research has presented contradictory results of either a positive or a negative correlation.

Finally, studies that compare a real situation with its VE *interactive* simulation are presented in detail. This set of studies commonly involves HMDs, exploring how relevant performance is compared, for instance, to more traditional displays. In order to assess the efficiency of a specific design or interface (for instance, head tracking vs. mouse) it seems that in most relevant studies the only measure at hand is the performance of a specific task. I feel that this strategy is limiting. Although human performance efficiency is crucial for successful VE applications, the complexity of the virtual interface demands that additional measures be employed. One of the most common tasks and one inherently linked with the sense of ‘space’ that a computer graphics world strives to achieve is spatial perception or spatial memory tasks. I will show that by employing approaches that have been adopted after decades of experimentation in the human perception area, one could reveal aspects of technology that could not be revealed by taking into account only the performance of a task. The next chapter will introduce this rationale and subsequently the methodology followed by the experimental work of this thesis.

Chapter 3

A Methodology based on Memory Semantics

Spatial perception is primarily linked with reasoning about three-dimensional ‘space’. Computer graphics strive to achieve a similar sense of space and action (reaching out, moving through) as in the real world by means of photorealism, high update rate or intuitive interfaces for VE applications. Spatial perception is closely connected with spatial memory [Bad97]. It is challenging to identify whether VE simulations and virtual interfaces have an effect on the cognitive, mental processes participants employ in order to achieve a spatial memory task in a VE in relation to the mental processes employed for the same task in the real world. Thus, it is significant to focus on the actual perceptual mechanisms that participants follow in addition to their performance of a task.

Memory, in the sense of ‘information’ for subsequent analysis, plays an important role in perceptual systems such as the visual, auditory, haptic and kinesthetic. The use of a single term for memory might initially suggest that memory is a unitary system, albeit a complicated one. Clearly though, it is not one but many [Bad97]. This chapter outlines the main characteristics of human memory concentrating on spatial perception and mechanisms of retrieval. A theory of memory awareness states introduced by Tulving [Tul85] is presented in detail. This theory focuses on ‘how’ humans retrieve information as opposed to ‘what’. Following an initial review of relevant spatial perception studies in VE literature, the theory of memory awareness states is adjusted into a ‘metric’ that could assess the simulation fidelity of a VE application compared to its real world counterpart from a *cognitive* rather than a task point of view. This metric could detect variation between the cognitive processes participants employ in order to accomplish a spatial memory task under conditions such as stereo or monocular visual displays, head-tracking or mouse input, especially when the actual performance of the task does not reveal any statistically significant differences across conditions.

3.1 Memory and Perception

Human Memory is a system for storing and retrieving information acquired through our senses [Bad97], [Rie97]. The briefest memory store lasts for only a fraction of a second. Such sensory memories are perhaps best considered as an integral part of the process of *perceiving*. Both vision and hearing, for instance, appear to have a temporary storage stage, which could be termed *short-term* auditory or visual memory and that could last for a few seconds. In addition to these, though, humans clearly retain *long-term* memory for sights and sounds. Similar systems exist in the case of other senses such as smell, taste and touch. In this section, theories of mental imagery and memory are reviewed in detail offering a complete illustration of the elements employed towards constructing a methodology which will aid VE systems’ simulation fidelity evaluation.

3.1.1 Mental Imagery

How many rooms exist in one’s house? What kind of clothes did one wear yesterday? These questions evoke *visual mental imagery*. In order to answer the first question, one would normally visualise his/her house and then ‘scan’ each room with an internal ‘eye’. Similarly, if asked how a sound differs from another sound, one would report ‘hearing’ the sounds in their mind’s ear. Such tasks *evoke auditory mental* *imagery*. Visual mental imagery is ‘seeing’ in the absence of the appropriate immediate sensory input. Auditory mental imagery is ‘hearing’ in the absence of the appropriate immediate sensory input. Imagery is distinct from perception that is defined as the registration of physically present stimuli [KBJ95]. Mental images, generally, are viewed as a form of mental representation generated internally without adequate external stimulus. These produce behavioural effects similar to those obtained with corresponding perceptual representations [Coo95]. However, demonstrating the relationship between mental images and percepts leaves open questions about the appropriate format for either sort of representation, at levels deeper than that of conscious experience.

The examples above illustrate that imagery plays an important role in memory and spatial reasoning, however, imagery also plays a role in abstract reasoning, skill learning and language comprehension. Even within a single sensory modality imagery is not a single undifferentiated ability, but rather, imagery involves a host of processes working together. The image of one’s house must be initially formed, it then must be scanned and also maintained while this is occurring. Moreover, in many situations an imaged object must be transformed. For example, when asked what can be found on the backside of an object, participants often report ‘mentally rotating’ the object in question and ‘looking’ at the relevant side of it.

Interest in mental imagery can be traced back to the time of Plato who thought that memories were based on images. According to his theory, memories are carved into the mind much like pictures can be carved on a wax tablet. Plato even took account of individual differences in terms, for instance, of the ‘purity’ of the wax and in the ease of carving figures into it. In more recent days, a lack of sophisticated methodologies for studying internal mental events followed a wide criticism on the value of imagery research. However, as the limitations of approaches perceiving the study of behaviour as the only scientific method became apparent, researchers started investigating internal events. Subsequent research showed that mental imagery interferes more with like-modality perception than with different modality perception, e.g. visual imagery interferes with visual perception more than with auditory perception and vice versa for auditory imagery [SF70]. Also, research showed that despite the fact that images are not actual objects that must obey the laws of physics, objects in mental images often behave like actual objects [SM71]. The difficulties in characterising the way images are represented and processed led researchers towards using neuropsychological data to inform theories of the structure of the processing system. Major areas of research for the study of mental imagery are learning and memory, perception and action, information processing and reasoning.

Learning and Memory

The ancient Greeks discovered that one's memory for a set of objects could be greatly enhanced if one visualised them interacting in some way. Once one has visualised a scene and encoded it into memory, it can then be recalled in an image. Much of the mnemonic power of imagery comes from its ability to represent associations between distinct objects. Even if people do not set out to store images intentionally, they will often use imagery later to recall information. Imagery is used to recall the shape, colour, size and texture of objects or spatial relations in scenes that cannot be inferred from facts associated strongly with them. For example, imagery is used to determine whether somebody wears a blue jacket but not whether the same person has two legs. Cooper [Coo95] states that in various studies, instructions to employ mental imagery as a mediator or mnemonic device could greatly enhance performance on standard tests of memory.

Perception and Action

Imagery draws on mechanisms used not only in perception but also in motor control. People sometimes speak of 'seeing' objects in visual images or 'hearing' sounds in auditory images. It is also possible that the motor system plays a significant role in imagery, specifically in 'image transformation' processes as noted above. Goldenberg et al. [GMN95] reported the case of a brain-damaged patient who had cortical blindness but still had visual mental imagery. This patient denied that she was blind and this belief could have been based on her confusing visual mental images for actual percepts. The authors note that this patient's imagery could have been triggered by tactile or **acoustic** perception. Once her vision recovered in the portion of the field that was registered by the preserved cortex, non-visual input no longer appeared to induce the illusion of seeing. These findings suggested that connections from the non-visual cortex play a complex role in evoking imagery.

Information Processing

Representations of images, in which information is stored, are processed in various ways. The system in which images are processed can be described in terms of these major abilities:

- Image inspection is the ability to extract information that is depicted in an image. In the course of inspecting imaged objects, participants of an experimental study report having to 'zoom in' to 'see' small details, and, in fact, participants require more time to inspect imaged objects of small sizes than objects imaged at larger sizes [Kos95].
- Image generation is the ability to retain sensory input and activate long-memory mechanisms. Many mental images are novel combinations of objects or characteristics that were encoded at different times and places. Many studies have shown that image generation involves serial processing and thus, the time to form images typically increases linearly for each additional part of the to-be-visualised object. The time to visualise a part often can be predicted by the order in which participants typically draw the parts.
- Image transformation is the ability to apply certain transformations to mental imagery. Image transformations often preserve the time course of the corresponding *actual* transformation.
- Image retention refers to the ability of maintaining an image for imagery tasks that require considerable time to complete.

Reasoning

Mental imagery is significant since this is a means by which information is learned, stored and retrieved. Consider how one decides what is the best route to get to work at rush hour in the morning or if a piece of furniture would fit in one's home. In both cases, imagery is used to carry out a kind of **'mental simulation'**. A major bottleneck in using mental imagery is the capacity of working memory; this capacity hinges on properties of a passive store and properties of active imagery operations. Mental imagery is also likely to be employed when for example, one visualises mentally a house following a verbal description of its arrangement, or even in abstract reasoning.

Generally, current distinctions between differences in visual/spatial representations emerging from research in cognitive science and neuroscience may apply to mental images as well. Visual imagery like visual representation or memory (analysed in the next section) may not prove to be a unitary concept. It should more accurately be viewed as referring to distinguishable subsystems specialised for performing particular aspects of cognitive tasks. Distinctions among subsystems of imagery specialised for particular tasks will become increasingly sophisticated, as researchers come to understand better the varieties of visual representation [Coo95].

3.1.2 Categories of Memory

According to Tulving, [Tul93], one could draw a classificatory scheme of five major memory and learning systems. These are:

Procedural Memory: This system is an *action* one. Its operations are expressed in behaviour and performance; they do not require conscious awareness of a kind that characterises other forms of memory. For instance, the ability to describe an action does not necessarily mean that one could actually perform it correctly.

Perceptual Representation: This is a non-conscious, *cognitive* form of learning that consists in the facilitation of the perceptual identification of words and objects by representing information about the form and structure but not the meaning and other associative properties of words and objects. It is considered a pre-semantic system and its operations can be carried out independently of the semantic and other higher memory systems.

Short-term memory: This *cognitive* system retains perceptual and conceptual information for a period of time measured in seconds after the input. It makes possible a conscious awareness of recently presented stimuli or of recently contemplated thoughts. It is dissociated from long-term (episodic and semantic) memory.

The following two long-term memory categories are considered of importance in this thesis:

Semantic Memory: The semantic *cognitive* memory system makes possible acquisition, retention and use of organised information in the broader sense; its principal function is cognitive modelling of the world. Semantic memory is necessary for the use of language. It is a ‘mental thesaurus’ including knowledge about words and verbal symbols, rules and algorithms for the manipulation of these symbols, concepts and relations.

Episodic Memory: The episodic *cognitive* memory system enables the individual to consciously remember personally experienced events embedded in a set of other personal happenings in subjective mental time. A perceptual event can be stored solely in terms of its perceptible properties or attributes. Every item in episodic memory represents information stored about the occurrence of an episode or event.

The semantic system permits the retrieval of information that was not directly stored in it and retrieval of information from the system leaves its contents unchanged although any act of retrieval constitutes an input into episodic memory. The semantic system is probably much less susceptible to involuntary transformation than the episodic system. Tulving [Tul93] argues that the semantic system may be quite independent of the episodic system in recording and maintaining information since identical storage consequences may be brought about by a great variety of input signals. He also states that the exercise of identifying various memory situations with episodic or semantic memory is not simple since many tasks contain both episodic and semantic features:

> 'The assignment of a task to one or the other category depends upon the kind of memory query addressed to the person, the exact nature of the information to be retrieved or the nature of the memory claim made about the retrieved information by the person retrieving it'.

While the specific form in which perceptual input is registered into the episodic memory can at times be strongly influenced by information in semantic memory -the phenomenon is referred to as encoding- it is also possible for the episodic system to operate relatively independently of the semantic system. Baddeley [Bad97] argues that a dichotomy in this sense might be an oversimplification, however, the lack of consistent terminology in this area probably reflects different views on the underlying processes.

It is useful to draw a distinction between consciousness and awareness. According to Tulving [Tul93] consciousness is determined by the properties of the individual's brain and general state at any given time. A given kind of consciousness determines what kinds of awareness or subjective experience the person *can* have; it provides one with a potential for particular kinds of awareness. To be aware of something means to have a particular subjective experience that is determined by both the current state of consciousness and the current stimulation from external and internal sources. Thus, awareness presumes consciousness, but consciousness does not imply awareness. Consciousness is a necessary but not a sufficient condition of awareness. Within a given level of awareness many particular kinds of subjective experiences may occur. Selective attention could be thought of as the primary process that determines the aspects of the stimulus situation of which the individual is aware; the direction of consciousness in the selection of 'contents' of awareness. Neither awareness nor consciousness should be considered as unitary systems.

It has been suggested that it may be useful to imagine a rough parallel between various forms of consciousness and different kinds of memory. The issue of conscious awareness becomes more central in the distinction between *explicit* and *implicit* retrieval [Tul93]. These labels refer to the presence or absence of the participant's awareness at the time of retrieval of the relation between present and past experience. In explicit retrieval, the awareness of the referential relation between the present and the past is usually referred to as 'remembering' or 'conscious recollection'. In implicit retrieval, a similar referential awareness is said to be absent; one's present cognitive activity has been influenced by past experience but is unaware of this fact. Despite the close relationship between explicit and episodic memory and implicit and semantic memory, the systems are not the same. One is a form of retrieval and one is a memory system.

3.1.3 The Remember/Know Paradigm

In this section, the main methodology employed in this thesis is going to be analysed. This methodology forms the core of the experimental design presented in Chapter 4 and 5. It has been adopted and adjusted in the context of three experimental studies presented in those chapters.

In the process of acquiring a new knowledge domain, visual or non-visual, information retained is open to a number of different states. 'Remembering' and 'Knowing' are two subjective states of awareness linked with memory recollections. Some elements of a learning experience or of a visual space may be 'remembered' linked to a specific recollection event and mental image or could just pop-out, thus, could be just ‘known’. Tulving [Tul85] introduced a distinction between *‘remember’* and *‘know’* responses and provided the first demonstration that these responses can be made in a memory test, item by item out of a set of memory recall questions, to report awareness states as well. He reported illustrative experiments in which participants were instructed to report their states of awareness at the time they recalled or recognised words they had previously encountered in a study list. If they remembered what they experienced at the time they encountered the word, they made a ‘remember’ response. If they were aware they had encountered the word in the study list but did not remember anything they experienced at that time, they expressed a ‘know’ response. The results indicated that participants could quite easily distinguish between experiences of remembering and knowing. According to Tulving, recollective experiences are the hallmark of the episodic memory system. Remembering refers to personal experiences of the past, those that are recreated with the awareness of re-living these events and experiences mentally. Knowing refers to those in which there is no awareness of re-living any particular events or experiences. Knowing also includes the general sense of familiarity about more abstract knowledge. ‘Know’ responses are the hallmark of the semantic memory system. Following this first major theory, episodic memory preserves the spatiotemporal properties of original experiences. Semantic memory does not preserve such information but, again, is a kind of mental thesaurus that retains conceptual representations.

Research conducted by John Gardiner and his associates, based on the distinction between ‘remembering’ and ‘knowing’ suggests that participants can arrive at knowledge about the ‘contents’ of past episodes even in word-list experiments in the laboratory. This could be achieved not only on the basis of episodic memory, but also, maybe less effectively on the basis of semantic memory [Gar2001]. There is some preliminary evidence that the distinction between ‘remembering’ and ‘knowing’ reflects a difference in *brain activity* at the time of encoding [Smi92]. Gregg & Gardiner [GG94], however, devised a procedure where studied words were presented visually at an extremely rapid rate in conjunction with a highly perceptual orientation task. Recognition performance was poor under these conditions and largely based on ‘knowing’. The test modality was subsequently manipulated after the initial visual presentation. Visual was compared with auditory presentation at the main study. It is reported that recognition performance was considerably better in the visual test where the test mode corresponded with the main study mode than it was in the auditory test where test and study modes differed. This modality effect occurred in ‘know’ responses, not in ‘remember’ responses. It is assumed following this study that recognition memory can be based largely on knowing, with little or no remembering. All that is necessary for encoding into the semantic system is some initial awareness of events. In contrast, encoding into episodic memory must depend on *greater conscious elaboration* of the events. Greg & Gardiner [GG94] also showed that estimates of the strength of the memory trace are greater when derived from remember plus know responses than when derived from only remember responses. Knowing, thus, reflects *an additional source of memory,* not merely a difference in response criteria.

Rajaram [Raj96] developed a second major theory. It was based on the suggestion that conceptual processes influence remembering. Knowing is influenced by perceptual processes. Recent evidence indicates that the relationship between the process distinction and the awareness distinction is more accurately regarded as orthogonal. Gardiner [Gar2001] states that

> ‘distinctiveness of processing, whether conceptual or perceptual, presumably engages a greater degree of conscious elaboration at study, and thereby enhances remembering. Knowing, in contrast, reflects fluency in processing either in conceptual or perceptual processes.’

This theory is complementary rather than alternative to the memory systems theory. Evidence that knowing is affected by conceptual as well as perceptual factors strengthens the semantic memory interpretation of this state of awareness. The two theories differ more in their explanations of knowing than in their explanations of remembering.

Although, ‘remember’ and ‘know’ awareness states have been controversially linked to episodic and semantic memory types, with ‘know’ responses more theoretically problematic, recent research emphasised that

> ‘they can be used without commitment to any theory, but simply to provide information on how various phenomena, including memory disorders, are characterised experientially’ [Gar2001].

In a relevant study, overall recognition performance in two groups of participants was very similar, however, the reported states of awareness differed markedly. Since one cannot make assumptions on what participants experience mentally from only their performance, there is no alternative to the use of subjective reports. Thus, additional information of awareness states, provides an invaluable input into ‘how’ participants remember as opposed to ‘what’. Subsequent research to [Tul85], summarised in [Gar2001], demonstrated that some variables affect one or the other of the two states of awareness, that some variables have opposing effects on them and that some variables have parallel effects on them. This finding indicates that the two states of awareness are functionally independent. According to Gardiner [Gar2001] this is a remarkable discovery and its full theoretical significance has yet to be appreciated.

3.1.4 Additions to the Remember/Know Paradigm

Conway et al. [CGPAC97] investigated the acquisition of knowledge by undergraduate students. For every memory recollection, the students were required to choose a memory awareness state between ‘remember’ and ‘know’ but also including ‘familiar’ and ‘guess’ states. The authors argued that ‘familiarity’ has at least two quite distinct meanings. It can be defined as the feeling that something has been encountered or experienced recently, although nothing about this recent occurrence can be remembered, it just feels familiar. ‘Know’ responses, on the other hand, represent highly familiar memory items that may come to mind without recollecting any particular encounter or any feeling of a recent encounter and cannot be placed. The items are just known. When ‘remember’ and ‘know’ are the only responses, participants of experimental studies could use ‘know’ responses to reflect strategies that do not involve any awareness. Allowing participants to report ‘familiar’ and ‘guess’ responses seems to be a good solution and evidence suggests that it is ‘guess’ responses rather than ‘know’ responses that reflect other strategies [Gar2001]. The study by Conway et al. [CGPAC97] that distinguished between a just ‘know’ response and a ‘familiarity’ response showed that these finer grained judgements can be dissociated from each other, just as different source memory judgements can. The authors suggested that when a new knowledge domain is to be acquired, memory is represented initially in an episodic way. As time goes by, the underlying representations may change such that they do not represent recollective experiences and are simply 'known' leading to a semantic representation. Good memory initially meant good episodic memory, which can presumably facilitate the development of more schematised conceptual knowledge in semantic memory. Such knowledge gives rise to knowing rather than knowing of recent, but not remembered encounters as measured by the familiarity responses. Again, there is little evidence that feelings of familiarity reflect the same memory system that supports highly familiar long-term knowledge as the semantic memory system. However, the remember-to-know shift in this study occurred in just the 'know' responses not in the 'familiarity' responses which like 'guess' responses were largely unchanged between the initial test and the subsequent retest.

Gardiner [Gar2001] concludes in his excellent review:

> 'The evidence reviewed strengthens considerably the case for arguing that psychology of memory should take on board subjective reports of conscious states and not just rely on more conventional measures of performance. This evidence has established that the essential subjectivity of remembering and knowing does not make reports of these states of awareness intractable to science'.

3.1.5 Memory Awareness States' Statistical Analysis

Koriat & Goldsmith [KG94] have drawn an important distinction between the amount or quantity remembered compared to the accuracy or quality of what is remembered. In the quantity analysis, memory awareness states data are represented as *a priori* or *prior probabilities.* Although this notation does not follow the general probability theory principles for 'prior' probabilities, it is going to be adopted as such in this thesis following the characterisations of Koriat & Goldsmith [KG94] as well as Conway et al. [CGPAC97]. Prior probabilities, here, are obtained by calculating the proportions of correct answers falling in each of the four memory awareness categories for every participant in a memory experimental study. In the accuracy analysis, correct recall scores are represented as *posteriori* or *posterior probabilities*. In order to calculate posterior probabilities, the proportion of correct answers from the total of answers given in each memory awareness category is computed individually for each participant.

For participant *n,*

x_{in} is the number of correct answers for the *i* awareness state,

x'_{in} is the number of incorrect answers for the *i* awareness state,

i = {remember, know, familiar, guess} = {1,2,3,4}

then,

P_{in} is the prior probability for awareness state *i* related to participant ***n*** (Equation 3.1),

$$P_{in} = \frac{x_{in}}{\sum_{i=1}^{4} x_{in}}$$

P'_{in} is the posterior probability for awareness state *i* related to participant ***n*** (Equation 3.2)**,**

$$P'_{in} = \frac{x_{in}}{x_{in} + x'_{in}}$$

Generally, prior probabilities answer the following question: Given that the response of a participant is correct, what is the probability that the participant has chosen a particular state on that question? Posterior probabilities, on the other hand, pose the following question: Given that a response of a participant was assigned to one of the four memory awareness response categories, what is the probability that the response is correct?

3.2 Spatial Memory and Perception in Computer Graphics

One of the major applications areas for VEs is simulation and training for implementations that are meant for skill transfer in the real world. Since spatial perception is closely linked with memory, there is a body of research that has explored issues of transfer of training, employing spatial memory tasks under distinct technological conditions. Spatial memory tasks are considered quite significant in that sense, since in order to accomplish any task using a VE implementation, one should realise the structure of a particular space first. Thus, the utility of VEs for any applications for which they are being proposed is predicated upon the accuracy of the spatial representation formed in the VE. In that sense, spatial memory tasks are considered essential and they are often employed in usability studies and benchmarking platforms that assess VE interface task efficiency. In this section, related spatial memory studies are reviewed focusing on the testing methodology and experimental design. Finally, an alternative methodology is conceptually introduced based on the remember/know paradigm described in the previous section.

3.2.1 Spatial Perception Studies in Computer Graphics Research

Arthur et al. [AHC97] examined participants' ability to reproduce a complex spatial layout of objects having experienced them previously under different viewing conditions. The layout consisted of nine common objects arranged on a flat plane. These objects could be viewed in a free binocular virtual condition, a free binocular real-world condition and in a static monocular view of the real world. The latter condition allowed the participants to observe the world from a single viewpoint. Performance was assessed using mapping accuracy (participants were asked to draw the position of certain objects in the room on a map after the VE exposure) and comparisons of relative inter-object distances. Mapping results showed a significant effect of viewing condition where, interestingly, the static monocular condition was superior to both the active virtual and real binocular conditions. The exposure time to the environment for each participant was not set. Participants were given as much time as they required to view the environment. Accurate spatial memory is a function of time [BT81], thus, this should be the same across conditions in spatial memory studies. There is no mention about the quality of rendering employed; this is a significant factor when a VE is compared to its real counterpart.

Two experiments were conducted by Bailey & Witmer [BW94] in order to investigate route and configuration knowledge acquisition in a VE. Route knowledge represents distances, orientation cues and landmark sequences within a navigational route. Configurational knowledge refers to how spaces are related to each other and is typically acquired through the use of a map. Sixty participants studied route directions and photographs of landmarks for a complex route either with or without a map, then were assigned to one of three rehearsal groups: A group that trained in the VE, a group that trained in the actual building and a symbolic rehearsal group that verbally rehearsed the route directions while referencing photographs of actual landmarks. Following training, all participants were tested in the actual building. Route knowledge was assessed in the actual building by recording the number of attempted wrong turns and route traversal time. Knowledge of the overall building configuration was measured using a target location triangulation technique. This technique involves a CRT-based task and a paper-based task that yielded accuracy, consistency and distance estimation error scores from triangulating four targets in the building from three sighting locations. Participants level of presence was also assessed using the Witmer & Singer questionnaire [WS98] as well as the level of simulator sickness induced using the SSQ questionnaire [KLBL93]. The results indicate that route knowledge can be acquired in a VE and that it transfers to the real world. Participants trained in the actual building made significantly fewer wrong turns than participants trained in the VE and symbolically and also their traversal time of the building was shorter. There were, however, no significant effects of training group on configurational knowledge. The different time that each participant spent in the training environment as well as the testing strategy causes doubts for the validity of these results. Since participants were monitored for traversal time, the configurational knowledge test could be biased. The authors of this paper also indicated a weak positive correlation of perceived presence with task performance as well as a negative correlation of presence with simulator sickness. Both results have been supported in recent literature [SSetal98] but they have also been contradicted as analysed in Chapter 2 of this thesis.

Waller et al. [WHK98] examined the variables that communicate the transfer of spatial knowledge and discuss the form and development of spatial representation in VE training. Six groups were trained in six different environments including no training, real world, map, VE desktop, VE immersive and VE long immersive. The groups were then asked to apply route and configurational knowledge in a real world maze environment. Results indicated that short periods of VE training didn't make a difference to map training, however, longer exposure to the VE environment was more effective than real world training. Longer exposure was not, however, incorporated across all conditions. It is not clear if better training was due to the VE technology or to the longer exposure no matter what the training medium consisted of.

Bliss et al. [BTG97] compared no training, blueprint training and VE training for fire fighters to navigate a rescue route in an unfamiliar building. After training, the authors monitored the total navigation time and number of wrong turns exhibited by fire fighters in the actual building. Participants were required to rescue a mock baby (life sized doll) following the specific training route. Measures of performance were compared among groups of 10 participants each (between subject design). The results indicated that fire fighters trained with the VE or blueprints performed a quicker and more accurate rescue than those without training. The speed and accuracy of rescue performance did not differ significantly between the VE and blueprint training groups. There could be some ambiguity, though, related to the training sessions employing the blueprints or the VE. The participants are assigned a task during training but this inherently involved varied training times.

Dinh et al. [DWH99] conducted a study to investigate the effects of tactile, olfactory, audio and visual sensory cues on participants' sense of presence and memory recall of the environment and the objects in that environment. The authors utilised two levels of visual detail by reducing texture resolution with or without ambient auditory stimulation, olfactory stimulation and tactile stimulation. The presence assessment was conducted with the use of a questionnaire on a 0 to 100 scale. The participants were requested to answer four spatial layout questions and five object location questions by selecting one of eight choices representing all the rooms in the structure and also 'Nowhere' and 'Do not remember' choices. Results indicated that perceived presence was significantly higher for the conditions with auditory and tactile cues. In terms of spatial layout recall, no significant main effect was revealed across conditions. This could be due to the small amount of spatial layout questions. In terms of object locations' recall, a significant main effect of condition was revealed. Accurate recall was higher for the participants who experienced tactile cues and olfactory cues in their environment. The number of participants in this study was quite high (322 participants).

Henry et al. [Hen92] measured the extent to which VEs succeed in providing accurate perceptions of the basic characteristics of architectural spaces. He conducted a series of experiments which evaluated three simulation conditions including monitor display condition, stereo HMD condition, non-head tracked, stereo HMD condition and head tracked against the real world setting (control condition). A size estimate task and an orientation task were employed while participants navigated the VE. Additionally, a spatial memory task was incorporated in the design requiring participants to draw a plan of the building after their exposure to the VE. The resulting sketches were rated according to participants' perception of the rooms relative to each other, the path of the visit and their ability to rank the spaces from smaller to largest. The overall results show that participants in all four conditions had a rather accurate cognitive map of the building. They also suggest that participants in the head tracked condition achieved the highest accuracy of drawing a map of the building. The authors do not mention if this result was significant across conditions. The number of six participants per condition in this study is small and also, the model used for the simulations was of low quality of rendering.

Arthur [Art2000] explored the effect of Field of View (FoV) on performance with HMDs. For his Ph.D thesis, he conducted experiments using a custom, wide FoV (176 degrees) HMD. He found that performance and participants' sense of presence was degraded even at the relatively high FoV of 112 degrees and further at 48 degrees. The experiments used a prototype tiled wide FoV HMD to measure performance in a VE with a custom, large area tracking. A relatively low number of five participants had to complete a number of tasks including a search task, a walking task, a distance estimation task and a spatial memory task. The spatial memory task involved certain circle-shaped objects appearing on the floor plan of a room, which were viewed under varied degrees of FoV. The participant was instructed to focus on the positioning of the objects. The memory task was competed on a monitor screen using a mouse by clicking and dragging the objects from the left side of the screen across the relevant positions. No main effect of FoV was identified for the spatial memory task and also participants' perceived presence was not higher under the widest FoV condition. The small number of objects (five) as well as participants (five) does not guarantee the validity of these results.

Billinghurst & Weghorst [BW95] investigated memory recall of participants navigating a virtual space. They assessed accurate recollection by asking participants to make a sketch of the space they had experienced. They correlated sketch results to a survey which included questions on a range of navigation, orientation, interaction, presence and interface questions. The study was designed to assess the validity of sketch maps as a tool for measuring topological knowledge of a VE. Following VE training, the participants were asked to complete the survey. Then, the participants experienced one of three different VEs for ten minutes and were requested to explore them in detail before producing a sketch map of the environment. A high positive correlation was revealed between subjective ratings of orientation and interaction and sketch map accuracy. The authors did not correlate sketch map accuracy which results in some ambiguity due to personalised style of drawing, to actual spatial questions related to elements of the VE space.

As demonstrated in this section, spatial studies in VE literature mainly focus on accurate spatial memory recall scores and incorporate a wide range of experimental designs and testing methodologies. These studies do not employ a specific framework in most cases. In the next section, a methodology based on the remember/know paradigm described earlier is going to be discussed. This methodology is going to be employed in the experimental work of this thesis.

3.2.2 A Model of Human Performance based on Memory Semantics

A central research issue for VE applications for training is how participants mentally represent an interactive computer graphics world and how their recognition and memory of such worlds correspond to actual conditions. For all the spatial memory studies mentioned above, it is apparent that accurate spatial orientation and spatial recall is perceived as a measure employed to assess the effectiveness of a particular interface or design. The means of presenting information to trainees range from traditional maps to VE displays. In some cases, presence and simulator sickness assessments are employed as well. A strong link between presence and task performance is not revealed and therefore, these assessments are not central in the evaluation of specific applications. Generally, the research presented in the previous section investigates the suitability of VE systems as effective training mediums in comparison to more traditional means of training such as maps, blueprints or even desktop monitors.

Experimental post exposure methodologies for spatial recall investigations range from questionnaires to asking participants to draw sketches and maps of a space they experienced or combine the latter with distance estimation tasks. A more direct way of testing requires participants to apply their knowledge of a space, acquired though training across varied technological conditions, so as to navigate effectively the real world space represented. Different allocations of time or, in some cases, no allocation at all, compromises the results [AHC97], [WHK98] as accurate memory recall is a function of exposure time [BT81]. Also, a small number of participants [Art2000] as well as bypassing the issue of the actual rendering quality, especially in comparisons with real-world situations [AHC97], [Hen92], [WS98] could also influence the experimental design. The effects remain unknown. These issues are generic to experimental design and they are not related to the actual experimental methodology. The most significant issue in the studies analysed in the previous section as well as in Chapter 2 is the testing methodology for memory recall, focused on the performance of a task.

Certain methodologies in traditional spatial memory research are actually centred on the scores of accurate recall. For instance, in a classic study by Brewer & Treyens [BT81] the authors examined the episodic place information retained after a brief exposure (35 seconds) to an experimental room. Participants were brought into what they thought was a graduate student's office and later were tested for memory of the room with either drawing recall, written recall or verbal recognition. This study is designed to investigate the use of schemata in memory performance. The basic assumption of the schema theories is that an individual's prior experience will influence how he or she perceives, understands and remembers new information. For instance, if a quick perceptual scan of a room indicates that there is a clock on the wall, hands and numbers are going to be automatically assigned to the clock when it is recalled, even if it didn't have any. This particular work was based on performance of the spatial memory task, however, it builds upon a strong theoretical foundation based on the schemata theory. Experimental studies are often designed without following a formal framework based on the fundamentals of the actual tasks involved. I feel that tasks such as spatial memory and distance estimation, widely employed in VE literature for evaluation purposes, have been theoretically documented in the fields of cognitive psychology, neuroscience and biology. This is work that should not be ignored. The field of VE is complex and is linked with various aspects of technological system design as well as human perception. Even if researchers cannot be experts on all fields, VE research should take input from already established theories already validated.

The remember/know methodology presented in a previous section in this chapter is a good example of a theory that has concerned researchers in the memory field for the last 30 years. It is a fascinating theory that allows experimentation to go beyond accurate memory recall scores following any testing methodology, uncovering the actual mental processes that participants employ to achieve a memory related task. It is established that VE performance measures need to focus on more than task performance to be effective due to the complex nature of simulating the human perceptual and motor mechanisms into a successful VE application. This theory, hence, could be perceived as a unified theory for any spatial memory task. It could be true that even if the environment and subsequently the spatial task's context in question varies, the underlying mechanisms identified through the remember/know paradigm are similar across environments under certain technological conditions and virtual interfaces. This hypothesis could satisfy the need of discovering a metric that is not going to be based on task performance as such, but it is going to be generically linked with it, theoretically and perceptually. As analysed in Chapter 2 of this thesis, the notion of presence perceived as a metric for evaluation of virtual interfaces was formed in order to assess designs and implementations without focusing on the specific performance of a task. The vision is that presence could be numerically constructed to fit in a parametric equation. By shifting parameters in that equation one could achieve the same sense of presence. This goal, however, is not yet fulfilled since there is not formal mathematical construct (and maybe there should not be) associated with presence. Existing research has not revealed a causal relationship or correlation between presence and task performance as demonstrated previously in this chapter. And why should such a relationship exist? A methodology such as the memory awareness states methodology expressed with the remember/know paradigm could be a valid approach towards linking task performance with cognitive processes.

In the following two chapters, three studies will be presented, the first perceived as a preliminary investigation and the second/third ones as the main studies. All studies incorporate a memory awareness states investigation, adjusting the extended remember/know theory in the context of a spatial memory task across varied displays and interfaces including traditional desktop monitors as well as HMDs. The displays could be head tracked or not, displaying mono or stereo imagery. The effect of the virtual interfaces in question on spatial memory awareness states is examined thoroughly using formal statistical methods based on prior and posterior probabilities. At the same time, assessments of perceived presence, simulator sickness and subjective impressions of lighting are incorporated in the design.

3.2.3 Proprioception Contributions to Navigation in Virtual Environments

In spatial memory studies, traditional input interfaces such as mouse-like interfaces are often compared to more ‘intuitive’ interfaces such as head tracking [WHK98], [PPW97]. Generally, navigation depends on realising self-position and orientation by piloting, path integration and orientation. Piloting relies on the observation of known landmarks and the ability to identify the spatial relationship between the landmarks and the observer. Path integration involves monitoring of the velocities or accelerations experienced while travelling. Integration of these cues will result in the navigator’s perception of current position relative to the starting point of the journey. *Proprioceptive* information reflects the movement of body parts relative to one another. It is necessary for co-ordinated bodily actions and is gained though mechanical receptors in joints and within the vestibular system but also through vision [BGG96]. The vestibular system is centred on the organ in the inner ear involved in the transduction of angular acceleration of the body into nerve impulses. An experimental study by Grant & Magee [GM98] investigated the contribution of inadequate proprioception to disorientation caused by immersive VEs towards transferring the spatial knowledge acquired to a real world task. Participants were provided with interfaces to a VE that either did (a walking interface) or did not (a joystick) afford proprioceptive feedback similar to that obtained during real walking. The two groups explored a large complex building using a low resolution HMD. Their navigational abilities (orientation and ability to find the shortest path to a given destination) within the actual building were compared with those of control groups. These studied a map, walked through the real building or received no prior training. Results showed that the walking interface conveyed no benefit on the orientation task performed during training in the VE but it did benefit participants when they tried to find objects in the real world. In another relevant study, Slater et al. [SUS95] used foot movements to toggle the participant's state between standing still and moving forward at a fixed velocity. This system proved to promote a higher sense of perceived presence. In a more recent study, Usoh et al. [UAWBSSB99] replicated the Slater et al. [SUS95] study adding real walking to the walking-in-place and the push-button-fly interface. This study confirmed the previous findings with subjective presence higher for real walking than walking-in-place involving, though, a weak overall effect of condition. Real walking was found to be significantly better than both walking-in-place and flying as a mode of locomotion. These studies did not examine spatial perception. Exploring the relationship between adequate proprioception and spatial memory is still an open research question. In this thesis, a strategy to monitor participants' navigation patterns is introduced in Chapter 5 across all conditions in an effort to identify relevant correlations of proprioceptive information with memory recall.

3.3 Summary

In this chapter, theories of mental imagery and human memory were initially introduced. Mental images, generally, are viewed as a form of mental representation, generated internally without adequate external stimulus that produce behavioural effect similar to those obtained with corresponding perceptual representations on tasks for which both forms of representation could be relevant [Coo95]. Mental imagery is linked with memory and strategies of retrieval. As with various perceptual systems, including mental imagery, memory is not a unitary system. Two significant memory subsystems, relevant to long-term memory are episodic and semantic memory. The semantic memory system is a 'mental thesaurus' including knowledge about words, verbal symbols and language. The episodic memory system enables the individual to consciously remember personally experienced events embedded in a set of other personal happenings in subjective mental time. Some elements of a learning experience or of a visual space may be 'remembered' linked to a specific recollection event and mental image related to episodic memory. Some elements could just pop-out, thus, could be just 'known' related to semantic memory. Tulving [Tul85] introduced a distinction between 'remember' and 'know' responses and provided the first demonstration that these responses can be made in a memory test, item by item out of a set of memory recall questions, to report awareness states as well. Recalls could also be reported as 'familiar' or even as 'guess', even if informed.

Spatial memory studies in computer graphics research are focused on accurate memory recall scores. Relevant research positively correlates the actual scores of accurate memory recall with the effectiveness of a specific VE implementation or virtual interface. This task dependent strategy, though, is limiting since tasks are always going to vary depending on the application. Although generalising results from specific experimentation is quite challenging, a strategy that is centrally focused on task performance for complex interfaces such as the virtual interfaces is not proven to be efficient with often contradictory results [SSetal98]. In the last section of this chapter, thus, I describe a methodology based on the remember/know paradigm of memory awareness states. Since accurate spatial perception is essential towards successful completion of any task, this methodology concentrates on the actual mental processes that participants employ in order to complete a spatial memory task together with accurate recollection scores. By employing methods that have been validated though decades of experimentation, adjusting them to fit the goals of computer graphics experimentation, computer graphics research applies informed practices for assessing the simulation fidelity of VE applications. It is fascinating to actually attempt to uncover the variations of perceptual mechanisms employed in the perception of synthetic spaces towards completion of tasks in an artificial world. These variations could differ depending on the display technology and interaction interface. It is also significant to examine how these processes would compare to the actual processes used in forming mental representations of spaces in the real world. Experimental studies which employ an adjusted and extended remember/know methodology in order to compare a real world situation to its interactive computer graphics simulation are presented in the following two chapters.

Chapter 4

A Preliminary Study

This chapter outlines the experimental methodology employed and the relevant results for the first, informally designed, preliminary study which compares spatial perception and memory recall in a VE displayed on a desktop monitor as well as on a HMD with its real situation counterpart [MC1999], [MC2000], [MCTH2000], [MC2001]. In this study, a methodology based on the remember/know paradigm presented in Chapter 3 is employed in the testing process towards identifying the *mental processes* participants followed for completing a memory task as well as the amount of their accurate recollections. The general goal is to identify variations of cognitive strategies related to their processes of retrieval (visual or not visual) as discussed in Chapter 3 when task performance across conditions does not differ. The actual task consists of two parts: non-visual information recall for participants experiencing a seminar-like situation and spatial recall of the environment where this experience was taking place. For the non-visual part of the task, an audio-only condition is also included in the experimental design. The computer graphics rendering of the real scene is retained non-photorealistic, e.g. flat-shaded rendering. This preliminary study was designed to acquire a basic set of data for the simplest rendering as well as HMD display (monocular, non-head tracked). This set of elements will be built up in the main studies of Chapter 5 to include photorealistic rendering, stereo graphics imagery and more complex virtual interfaces such as head tracking. Participants are required, here, to complete a memory task and provide self-reports of their level of perceived presence and simulator sickness, the latter for the HMD condition.

4.1 Experimental Methodology

In this section, the experimental design is going to be described in detail. Initially, two small pilot studies are going to be presented with a summary of the main lessons learnt. Subsequently, experimental design issues such as participants, apparatus and materials including assessment strategies are going to be discussed. Since this is an initial exploratory study which aims to investigate the feasibility of the awareness states methodology, the experimental design is not strict. A set of issues arising from the preliminary study is going to be discussed at the end of Chapter 4. A formal experimental design is going to be employed for the two major studies presented in Chapter 5, specifically addressing those issues.

4.1.1 Pilot Studies

Pilot trials are common and useful in human centred experimentation. Experimental procedures such as a questionnaires, instructions and methods may be tested on a small sample of participants in order to highlight ambiguities for which adjustments can be made before the actual data gathering process begins. Two small pilot studies were designed for the preliminary study described in this chapter. The first was conducted in order to assess participants' comprehension of the memory testing instructions employing the remember/know paradigm. Written instructions should be clear and uniform across conditions. The second pilot study was conducted in order to finalise the structure of the presence questionnaire which would be administered after the completion of the memory test and also try out the final version of the computer graphics application. The questions from the Slater et al. [SSMM98] questionnaire were adopted, however, certain adjustments were made in terms of the actual wording for the visual and audio experimental conditions to be accommodated in the context of the application. Additionally, the format of the questionnaire (not its actual content) was altered from the original one [SSMM98]. A line was incorporated with distinct separations of levels for the Likert scale instead of the original box-like structure, for ease of use. In terms of the computer graphics application, navigation was restricted towards simulation of head movements. The preliminary study is described in detail as follows.

4.1.2 Methods

In this preliminary study, a methodology for simulation fidelity evaluation of VEs centred on a validated theory of memory recall awareness states (the remember/know paradigm) is presented. The actual task consists of two parts: non-visual information recall for participants experiencing a seminar-like situation and spatial recall of the environment where this experience was taking place. The study investigates how exposure to a computer-generated replica of the environment, displayed on a typical desktop display and a HMD would compare to exposure to the same environment and memory recall task in the real world from a cognitive rather than a task point of view. Subjective measures such as memory awareness states selection and perceived presence assessments are incorporated together with objective measures of memory recall, in a comparative study of a VE against the real world. The remember/know paradigm focuses on the actual mental cognitive processes that participants employ in order to complete a memory task rather than on the actual scores of accurate memory recall, as discussed in Chapter 3. The resultant accurate seminar and spatial memory recall scores and awareness states as well as participants' sense of presence is compared with that obtained from an analogous experiment in the actual real world space. The extent to which judgements of memory recall, memory awareness states and presence in the physical and VE are similar provides a measure for the fidelity of the simulation in question.

Four groups of 18 participants were recruited to participate in this study from the student population of the University of Bristol and Hewlett Packard Laboratories in Bristol, UK. 80% of the subjects from each group were male. All use computers frequently in their daily activities. Participants were randomly assigned to each group.

A between-subject design was utilised balancing groups for age and gender. Participants in all conditions were informed that they could withdraw from participation at any time during the experiments and they were naïve as to the purpose of the experiment. They were also asked if they had any knowledge relevant to the historic topic of the seminar and if they did, they were excluded. Participants had either normal or corrected-to-normal vision. According to the group they were assigned to, participants completed the same memory task, in one of the following conditions:

1) In reality, attending to a 15-minute seminar in a seminar room in the University of Bristol; referred to as the **real-world condition**.
2) Using a computer graphics simulation of the real world space with the real-world audio on a desktop monitor; referred to as the **desktop condition**.
3) Using the same application on a monocular, non-head tracked HMD with the real-world audio and with a mouse for navigation; referred to as the **HMD condition**.
4) Listening to the audio recorded during the real-world condition and completing the non-visual part of the task; referred to as the **audio-only condition**.

This study is based on the assumption that a 3D desktop display is less immersive than a HMD. Audio used in one condition as the only experimental sensory stimulus is perceived, in this study, as the least technologically 'immersive' condition.

Prediction

Presence and task performance were predicted to be significantly higher in the real-world condition relative to the desktop, the HMD and the audio-only conditions, thus offering a high benchmark. The main scope of this study was to show that by incorporating cognitive measures *together with* task performance measures, variations of participants' mental processes for memory recall would be revealed across conditions.

The Real Situation

The first group of 18 participants attended a seminar presentation that took place in a specific seminar room in the University of Bristol (Figure 4.1).

Figure 4.1: The real seminar room and the computer graphics environment.

	FoV (Field-of-View)	Resolution	Input Device
Desktop Monitor	38 degrees approx. hor.	1152*864	Mouse
HMD	30 degrees hor.	XGA(1024*764)	Mouse

Table 4.1: Technical characteristic comparison between the desktop monitor and the HMD.

The seminar's duration was 15 minutes. The historic content was chosen as none of the participants had any prior knowledge on this matter. The lecturer utilised 12 slides on an overhead projector. The seminar was digitally video recorded using a digital video camera on a tripod. Subsequently, the audio was extracted (16-bit stereo, 44kHz) in order to be incorporated in the computer graphics application for the desktop and HMD conditions.

The Graphical Simulation

The seminar room was modelled using the 3D Studio MAX [3ds2001] modelling package and converted to VRML (Figure 4.1) [VRML97]. The geometry in the real room was measured using a regular tape measure with accuracy of the order of one centimetre. The audio extracted during the real world seminar was incorporated in the computer graphics application. The application included a slide-show synchronised with the audio at the exact timings that the lecturer manipulated the slides in the real seminar. A static billboard with a texture displaying the lecturer (who was always facing the camera) was included in the application. The model was rendered flat-shaded and the application had an average update rate of 45 frames per second for both the desktop and HMD condition. The input device for navigation was a normal mouse in both the desktop and HMD conditions.

Figure 4.2: Desktop application.

The second group of 18 participants used the desktop application which included the audio recorded from the real seminar for the specified duration of the lecture (15 minutes) and their navigation tendencies were informally monitored (desktop condition). The application was displayed on a 21-inch typical desktop monitor (Figure 4.2). The FoV was calculated in relation to the distance of the participant from the display. A third group of 18 participants used the same application displayed on the HMD (HMD condition). A fourth set of 18 participants just listened to the audio recorded during the real seminar and completed the part of the memory task related to the seminar information. Obviously, the spatial perception task was not completed since there was no visual stimulus for this group (audio-only condition).

The HMD employed is a Hewlett Packard Laboratories working prototype and, thus, is not a commercial product available in the market. It is described as an ‘eye-glass’ display which features two micro-displays and appropriate optics, one for each eye. Both eyes are presented with the same image allowing for monocular imagery. Eyeglass displays allow for periphery vision and tend to be smaller and much lighter than fully-fledged HMDs as shown in figures 4.3, 4.4. The resolution of the desktop monitor employed in the desktop condition was kept at 1152*864; respectively, the resolution of the HMD was 1024*764. This small difference of FoV and resolution between the desktop monitor (38 degrees horizontal) and HMD (30 degrees horizontal) was considered minimal since this study was preliminary (Table 4.1).

Figure 4.3: The Hewlett Packard Laboratories HMD prototype.

Figure 4.4: Trying out the Hewlett Packard Laboratories HMD prototype.

Figure 4.5: Experimental set-up for the HMD condition.

The real world was perceived as a control condition so the FoV of the participants in the real-world condition was not restricted in this study. Participants in the desktop and HMD condition were able to explore the room from a steady viewpoint, approximately placed in the centre of the room. They had the ability to rotate on a full circle, horizontally, as well as on a half circle vertically, approximately emulating the movement of the head, using a common mouse (Figure 4.5). The experimental room was not darkened and participants in all conditions utilising computer graphics imagery were aware of their surroundings.

4.1.3 Materials

The four groups of participants were asked to fill in the same set of questionnaires after exposure. This included the memory task and memory awareness states questionnaire and the presence questionnaire [SSMM98] with the addition of the SSQ questionnaire [KLBL93] for the HMD condition. These materials can be found in Appendix A1.

Memory recall task

The questionnaire relevant to the memory task was designed to test the participants' accurate memory recall of the information communicated in the seminar and their spatial awareness of the environment. Overall, there were twenty-two questions. Sixteen questions were related to the actual factual information communicated in the seminar. The same set was incorporated in the real-world, desktop, HMD and audio-only conditions. Six questions were relevant to the environment where the seminar took place; these were incorporated in the real-world, desktop and HMD conditions. The correct answers for nine of the questions related to the seminar were included in the slideshow and were also mentioned by the lecturer. The remaining seven were only mentioned verbally and not included in the actual slide show.

Each memory recall question had four possible answers and it included a confidence measure with five possible states: No confidence, Low confidence, Moderate confidence, Confident, Certain. Most importantly, it also included an awareness state measure with four possibilities: Remember, Know, Familiar and Guess. Participants were required to select the correct answer for each question according to their recollection, select a confidence level and also report on their strategy of retrieval as expressed by one out of the four awareness states. This is based on the extended remember/know paradigm described in Chapter 3. Prior to filling out the core of the questionnaire, participants were given instructions that were designed to explain what each of the memory awareness states depicted as follows:

- You remembered a specific episode or image from the seminar. In this case you might have images and feelings in mind relating to the recalled information. Perhaps you virtually ‘hear’ again or ‘see’ again the lecturer presenting some item of information or remember visually the specific slide that information was included into. Answers such as these are called REMEMBER answers.
- You might just ‘know’ the correct answer and the alternative you have selected just ‘stood out’ from the choices available. In this case you would not recall a specific episode and instead you would simply know the answer. Answers with this basis are called KNOW answers.
- It may be, however, that you did not remember a specific instance, nor do you know the answer. Nevertheless, the alternative you have selected may seem or feel more familiar than any of the other alternatives. Answers made on this basis are called FAMILIAR answers.
- You may not have remembered, known, or felt that the choice you selected have been familiar. In which case you may have made a guess, possibly an informed guess, e.g. some of the choices look unlikely for other reasons so you have selected the one that looks least unlikely. This is called a GUESS answer.

Presence

The second questionnaire was designed to measure participants’ level of perceived presence on a Likert 7-point scale. The questions used in the Slater et al. study [SSMM98] were adopted and the questionnaire was administered in all conditions including the real-world one. This particular set of questions is exploring various aspects of the concept of presence itself and is not relevant to the technology or interface used for the application. For example, issues investigated were the dominance of the virtual world over the real one, the sense of visiting a ‘place’ versus viewing a scene or listening to a sound and the level that the memory of the experiment resembled everyday memories as presented in Chapter 2. Only questions of this nature could be applied to all four conditions without any per-condition tailoring. The questionnaire included additional questions regarding gender, ratings of background sounds, profession, level of computer-related expertise and level of losing track of time.

SSQ questionnaire

The widely used Simulator Sickness questionnaire (SSQ) was administered following participants' exposure to the VE for the HMD condition only. 16 symptoms were employed indicated in the Kennedy et al. study [KLBL93]. The questionnaire design is based on three components: Nausea, Oculomotor problems and Disorientation. Participants report the degree to which they experience each of the above symptoms as one of 'none', 'slight', 'moderate' and 'severe'. These are scored respectively as 0,1,2,3 and final scores are calculated as discussed in Chapter 2.

4.2 Results and Discussion of Preliminary Study

In this section, the results of the preliminary study are presented, separately for memory recall and awareness states, presence and simulator sickness. Certain correlations between different data sets are reported.

The memory recall task incorporated two sets of questions. The first set was relevant to the actual seminar information communicated by audio and visual slides and the second set was relevant to the participants' spatial perception, thus it included questions related to the seminar room itself. As mentioned above, every question had four possible answers, a confidence 5-point scale and a choice between four memory awareness states including 'remember', 'know', 'familiar' and 'guess'. The memory recall scores as well as the confidence scores were analysed using **AN**alysis of **VA**riance (ANOVA) [Coo99]. ANOVA is a powerful set of procedures used for testing significance where two or more conditions are used. Significance decisions involve rejecting or retaining the null hypothesis (which claims that groups are identical). The null hypothesis is rejected when the probability that a result occurring under it is less than .05. Generally, ANOVA compares the variance of the sample means (between groups variance) with the within groups variance. Within groups variance is calculated by taking the average of the variances within each sample around its mean.

4.2.1 Task Performance: Memory Recall for the *Non-Visual* Memory Task

Figure 4.6 shows the mean accurate recall scores (amount of accurate answers related to the seminar out of the sixteen questions) and confidence levels (5-point scale) for the task relevant to the seminar. A Post-Hoc Scheffé test was employed for multiple comparisons between the four conditions [Coo99]. Once a significant difference is determined among means, post-hoc range tests and pairwise multiple comparisons can determine which means differ. Range tests identify homogeneous subsets of means that are not different from each other. Pairwise multiple comparisons test the differences between each pair of means, and indicate significantly different group means at an alpha level of 0.05. The significance level of the Scheffé test is designed to allow all possible linear combinations of group means to be tested, not just pairwise comparisons available in this feature. The result is that the Scheffé test is often more conservative than other tests which means that a larger difference between means is required for significance.

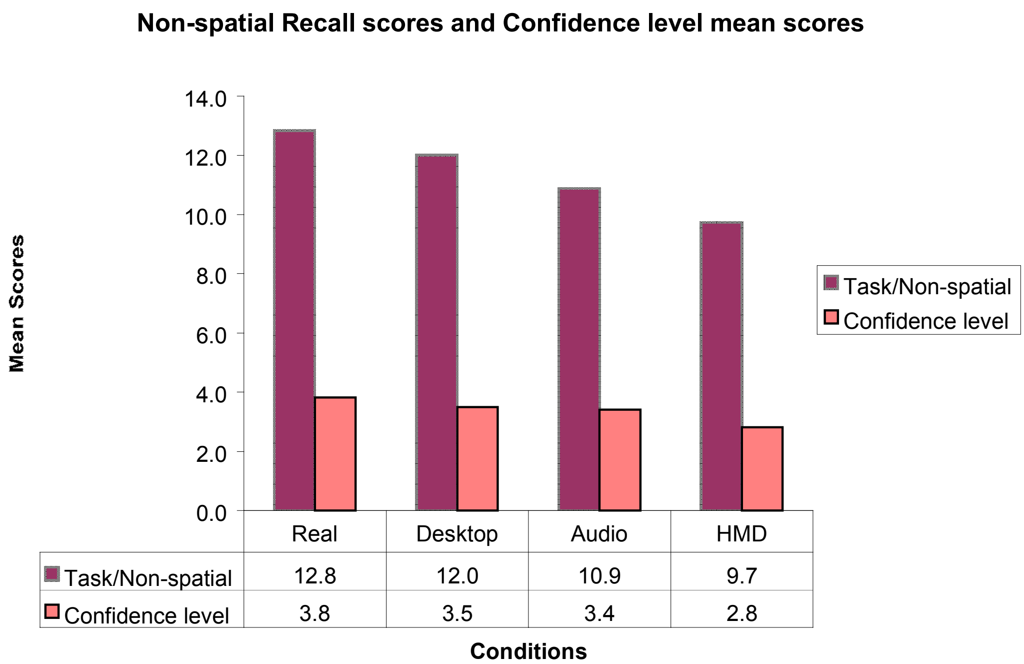
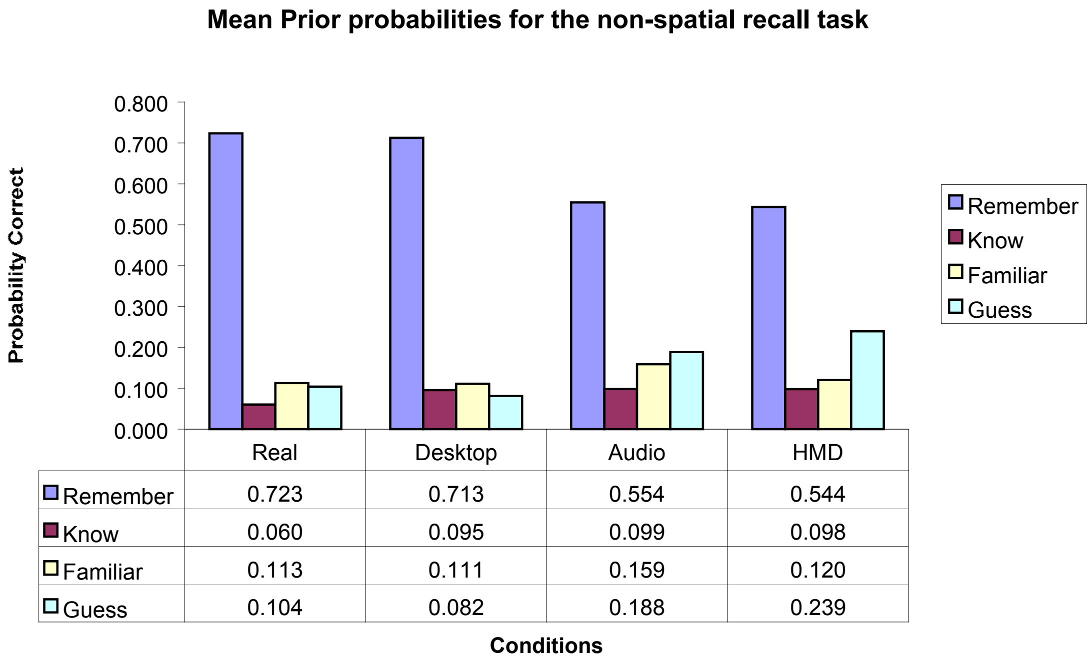


Figure 4.6: Mean accurate recall scores (correct answers out of sixteen questions) for the memory recall of the seminar information and relevant confidence scores.

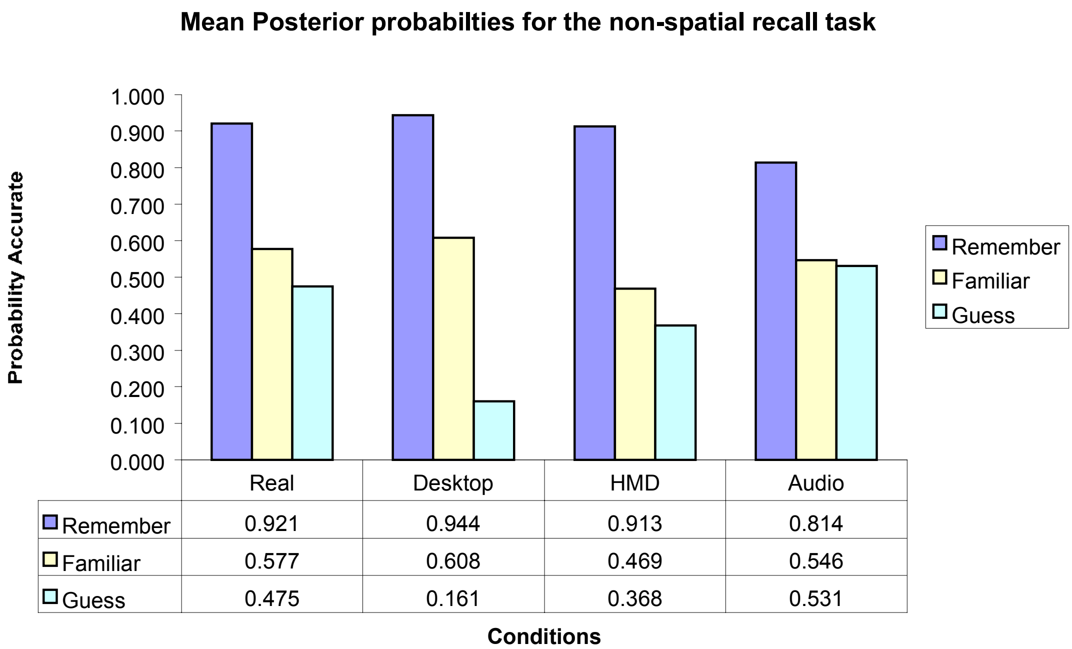
There was an overall significant effect of condition, $F(3,71)=6.590$, $p<0.05$, for the accurate memory recall of the non-visual part of the task. This equation can be read as follows, the F statistic equals 6.590, with 3 degrees of freedom (4 conditions) and 72 participants (between subjects' design). The p value indicates the probability that these differences occur by chance. The result, here, means that there are statistically reliable differences between the conditions. The F ratio is calculated by dividing the between group variance with the within group variance. If means differ among themselves far more than participants differ within groups, then the F ratio will be higher than 1 to a significant extent and this will be crucial for statistical significance. Memory recall for this portion of the task was significantly higher for the real-world condition compared to the HMD condition ($p<0.001$) and also compared to the audio-only condition ($p<0.05$) but not compared to the desktop condition. No significant differences of accurate recall were revealed between the audio-only condition and either the desktop or HMD conditions. However, the accurate recall scores for the desktop condition compared to the audio-only condition were significantly higher for those questions that had their answers written on the slides and were also communicated via the audio ($p<0.05$). This fact shows that relevant visual stimuli enhanced task performance under these conditions. The post-hoc test also revealed an increase in recall for the desktop relative to the HMD condition, an effect that approached significance ($p<0.06$).

Confidence scores followed approximately the same pattern as the memory recall scores. There was a reliable effect of condition upon confidence, $F(3,71)=8.582$, $p<0.01$. Confidence levels for the real-world condition were higher and approaching significance compared to the audio-only condition ($p<0.08$). The confidence scores in the real-world condition were significantly higher relative to the HMD condition ($p<0.01$) and also significantly higher for the desktop compared with the HMD condition ($p<0.05$). Although there was no statistical difference for accurate memory recall between the HMD and audio-only condition, the confidence scores for the audio-only condition were significantly higher compared to the HMD condition. Participants in the HMD condition achieved the lowest memory recall scores and also confidence scores, even using a light HMD such as the Hewlett Packard Laboratories prototype.



	Real	Desktop	Audio	HMD
Remember	0.723	0.713	0.554	0.544
Know	0.060	0.095	0.099	0.098
Familiar	0.113	0.111	0.159	0.120
Guess	0.104	0.082	0.188	0.239

Figure 4.7: Mean Prior probabilities by memory awareness state for the non-spatial task.



	Real	Desktop	HMD	Audio
Remember	0.921	0.944	0.913	0.814
Familiar	0.577	0.608	0.469	0.546
Guess	0.475	0.161	0.368	0.531

Figure 4.8: Mean Posterior probabilities by memory awareness state for the non-spatial task

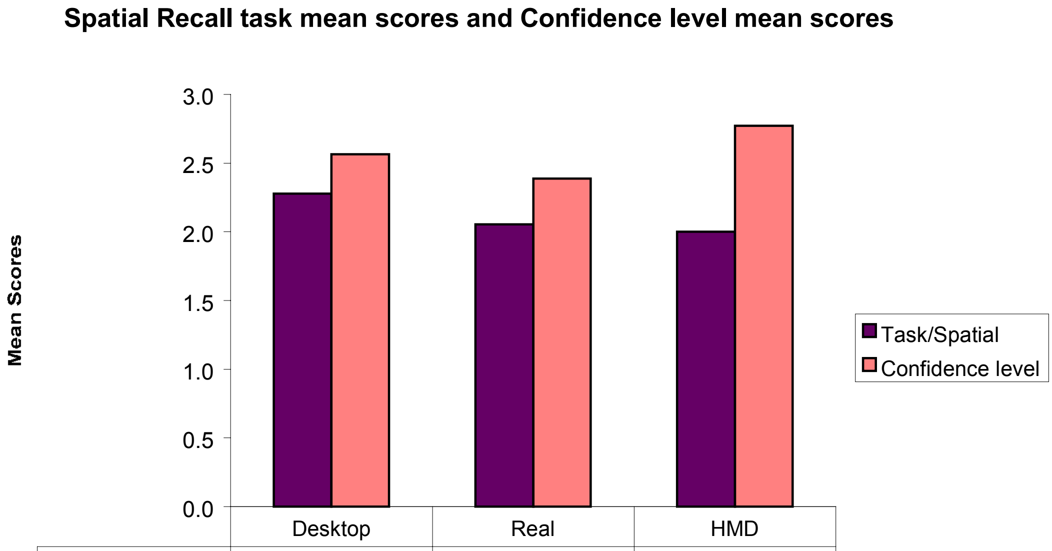
Figure 4.7 shows the mean prior probabilities related to the four awareness states across conditions for the seminar (non-spatial) recall portion of the task. Prior probabilities represent the probability that a correct memory recall response falls under a specific memory awareness state and indicate the proportion of correct answers for each of the memory awareness states. Hence, they represent the probability that a correct answer will be ‘attached’ to a specific state. If a participant made an accurate recollection, this is the probability that he/she would have chosen the specified state at the time of his/her recollection.

The memory awareness states’ data were analysed using ANOVA and Tukey’s Honestly Significance Post-Hoc Tests (HSD). HSD makes all pairwise comparisons between groups and sets the experimentwise error rate to the error rate for the collection for all pairwise comparisons. A significant main effect of condition upon the ‘remember’ awareness state, $F(3,71)=4.059$, $p<0.05$, and ‘guess’ awareness state, $F(3.71)=4.587$, $p<0.01$, was revealed. In particular, there is a tendency towards significance of a higher probability of correct responses under the ‘remember’ awareness state for the real-world condition compared to both the HMD ($p<0.08$) and the audio-only conditions ($p<0.06$). There is also a tendency towards significance of a higher probability of correct responses under the ‘remember’ awareness state for the desktop condition compared to the HMD ($p<0.1$) and the audio-only conditions ($p<0.08$). Additionally, a significantly higher probability of correct responses under the ‘guess’ awareness state for the HMD condition was revealed compared with the real-world ($p<0.05$) and the desktop condition ($p<0.05$) with a tendency towards significance for a higher probability of correct ‘guess’ responses for the audio-only condition compared with the desktop condition ($p<0.081$). It is noted that participants in the HMD condition had the lowest proportion of correct ‘remember’ responses and the highest proportion of correct ‘guess’ responses. The low confidence levels of participants in this condition verify this result. The real-world condition and the desktop condition revealed an equivalent proportion of correct ‘remember’ responses. The mental strategies followed as shown by the distribution of prior probabilities were similar for the real-world and the desktop condition. In that sense, for this part of the task, the desktop condition achieved a high level of simulation fidelity.

Figure 4.8 shows the mean posterior probabilities for the awareness states excluding the ‘know’ state due to a limited amount of answers assigned to that category across the four conditions. Posterior probabilities represent the probability that a memory recall response assigned to each of the memory awareness states is accurate. Calculating the posterior probabilities was rather problematic since the way these probabilities are calculated, according to Equation 3.2, does not guarantee that the denominator is not zero. This results in a number of posterior probabilities not being computed. The means on Figure 4.8 were based on the ones that could be calculated, thus, these means do not include all the participants and could only indicate a tendency. For posterior probabilities an ANOVA was not applied for this reason. For ANOVA statistics to be valid, a large number of memory recall questions need to be included in the experimental design to minimise the cases where the sum of the correct and incorrect responses under each memory awareness category is zero (no answers assigned to an awareness state). The mean posterior probability results show a higher probability for ‘remember’ responses to be accurate on all conditions. Also, a higher probability that ‘guess’ responses would be accurate was revealed for the real-world, HMD and the audio-only condition compared with the desktop condition. Statistical significance, however, across the experimental conditions for the posterior probabilities cannot be computed.

4.2.2 Task Performance: Memory Recall for the *Spatial* Memory Task

Figure 4.9 shows the mean accurate scores (correct answers out of six questions) and confidence levels levels (5-point scale) for the spatial awareness part of the task (Appendix A1). The questions related to this part of the task were centred on recollections of objects (or colours of objects) in the room where the seminar was taking place. An ANOVA did not reveal an overall effect of condition for this part of the memory task as well as for the confidence levels reported. No significant differences were identified across conditions.



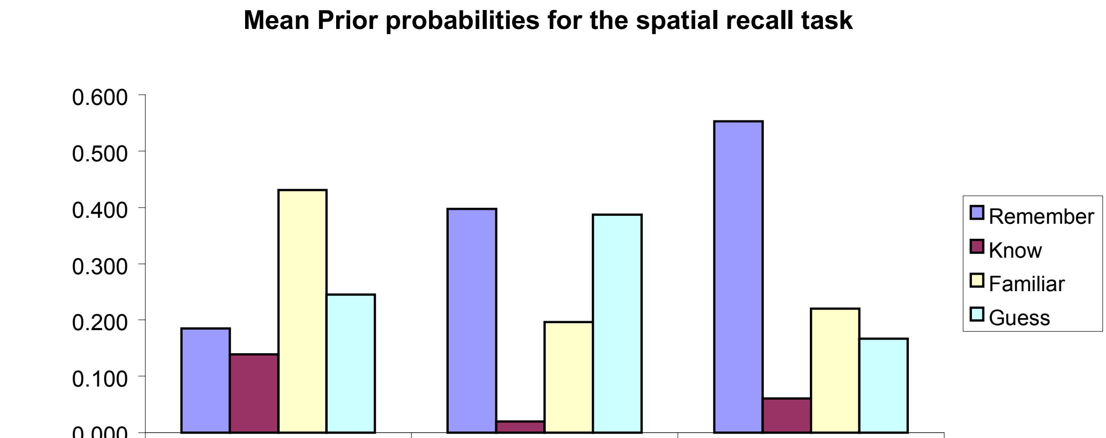
	Desktop	Real	HMD
Task/Spatial	2.3	2.1	2.0
Confidence level	2.6	2.4	2.8

Figure 4.9: Mean scores for the spatial memory recall task (correct answers out of six questions) and relevant confidence scores (5-point scale).

An ANOVA applied to the mean prior probabilities (Figure 4.10) for the spatial memory task showed a significant overall effect for the 'remember' awareness state, $F(2,53)=4.404$, $p<0.05$ and also for the 'familiar' awareness state, $F(2,53)=4.017$, $p<0.05$. This is a significant result as it sustains the original rationale behind employing the memory semantics methodology based on the remember/know paradigm. The actual methodology employed focusing on the actual mental processes of recollection did reveal statistically significant variations when task performance (memory recall) failed to do so. In particular, the probability that the correct responses would fall under the 'remember' state was significantly higher for the HMD condition compared with the real-world condition ($p<0.05$) but not compared to the desktop condition. In contrast, the probability that correct responses would fall under the 'familiar' state was significantly higher for the real-world condition compared to the HMD condition ($p<0.05$), also revealing a tendency towards higher significance relative to the desktop condition ($p<0.08$). The proportion of correct answers under the remember and know awareness states expressed by the prior probabilities also shows a higher level of confidence for the HMD condition in terms of participants' recollections. Although the actual confidence scores do not sustain this assumption, strategies of retrieval incorporate level of confidence information. The ‘familiar’ and ‘guess’ awareness states according to their definitions induce an amount of uncertainty in terms of the actual recollections assigned to those states.

Figure 4.11 shows the mean posterior probabilities which follow a similar pattern. The probability that ‘remember’ responses would be accurate was higher for the HMD condition than for either the desktop or the real-world conditions. An ANOVA was not conducted for these results for the same reason as for the results concerning the non-visual part of the task. For certain participants, there were no responses for correct and incorrect answers under the ‘know’ awareness state, thus, creating a problem in the calculations. The averages that are reported here do not, therefore, include all the participants and rather indicate a tendency.

As described in Chapter 3, the ‘remember’ awareness state is linked with episodic memory. ‘Remembering’ is defined as a state in which mental images relating to a past event or space come to mind while in the process of recall. Although there was not a statistically significant effect for the spatial memory recall scores across conditions, variations in the cognitive mental processes participants follow to actually process their recollections were revealed. This could be due to the novelty of the experience especially in the case of the HMD condition or to characteristics of the technology itself. A larger and more controlled study following a formal experimental design, which would focus on a visual task, should reveal if the above results could be generalised. This study, however, is an initial effort to investigate ‘how’ humans achieve a memory task in a VE from a cognitive point of view rather than a task performance point of view. These strategies could form a base for fidelity assessments of simulations.



	Real	Desktop	HMD
Remember	0.185	0.397	0.553
Know	0.139	0.020	0.061
Familiar	0.431	0.196	0.220
Guess	0.245	0.387	0.167

Figure 4.10: Mean Prior probabilities for the spatial recall task.

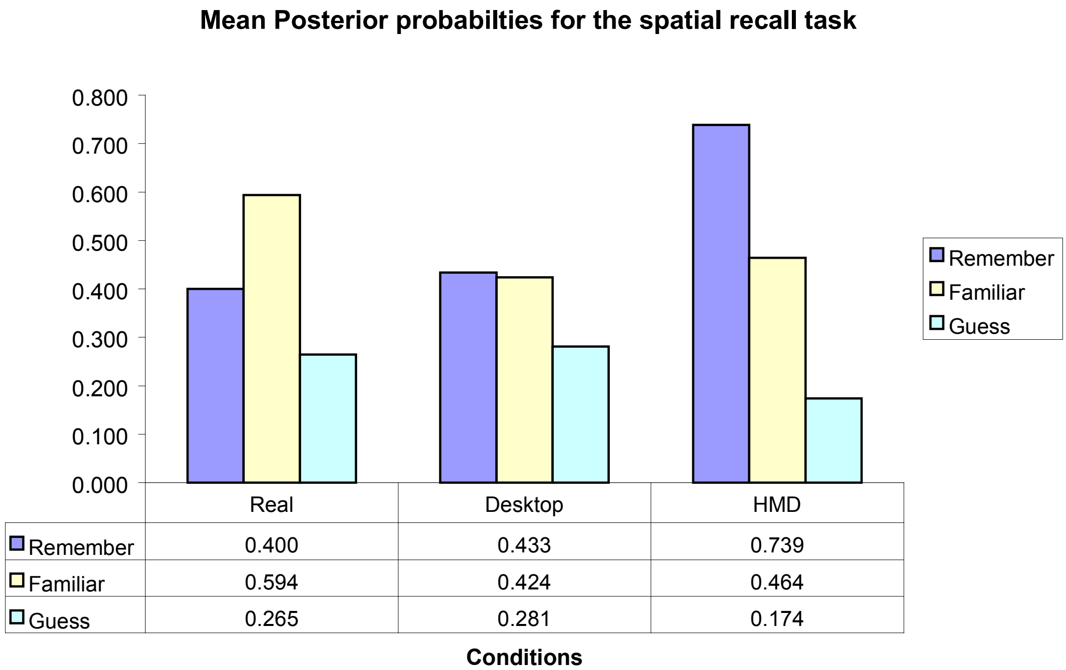


Figure 4.11: Mean Posterior probabilities for the spatial recall task across conditions.

4.2.3 Presence Results

The presence questionnaire was administered to the four groups assigned to the real, the desktop, the HMD and the audio-only conditions after the memory task was completed. Generally, there are various conceptual arguments concerning the definition of the notion of presence itself and its relevance to the real world. Researchers argue that the definition of presence is not related to the real world experiences but only to ‘mediated’ ones. It is also argued, though, that presence is after all defined as the ‘sense of being’ in a synthetic space similarly to a real world space and therefore, real world perceptions should be taken into account. In this study, the presence questionnaire was administered to the group assigned to the real-world condition in order to apply a uniform experimental design as well as attempt to validate it as a measuring ‘device’ for presence. This particular questionnaire has been developed through years of experimentation in UCL, London [SSMM98] and it was expected to ‘pick up’ the difference between the real world and the technological conditions. A recent study, however, by the researchers that developed it indicated a weak tendency towards significance when they administered the questionnaire in similar conditions [UCAS2000].

Presence data (7-point scale) were analysed using a comparison of means before carrying out an ANOVA. A Post-Hoc Scheffé test was also employed in order to obtain multiple comparisons' results between the four conditions. In addition to this generic analysis and to avoid the theoretical problem of ordinal data, a binomial regression analysis was used for the presence response based on the count of high scores (6 and 7) out of six presence questions. Generally, the gaps between the Likert-scale values are not necessarily equally spaced. The points on the scale have order (1 lower than 3) but not necessarily arithmetic properties but *ordinal* properties. The statistical analysis procedure presented in the Slater et al. study [SSMM98] was adopted. The binomial regression was accomplished by assigning 0 if the count of high scores was 0-2 and 1 if the count of high scores was 3-6. Binomial regression, generally, shows the probability of falling under one of the 0 or 1 binomial distributions for each group (condition). This method verified the results regarding statistical differences acquired from the generic ANOVA statistics described here.

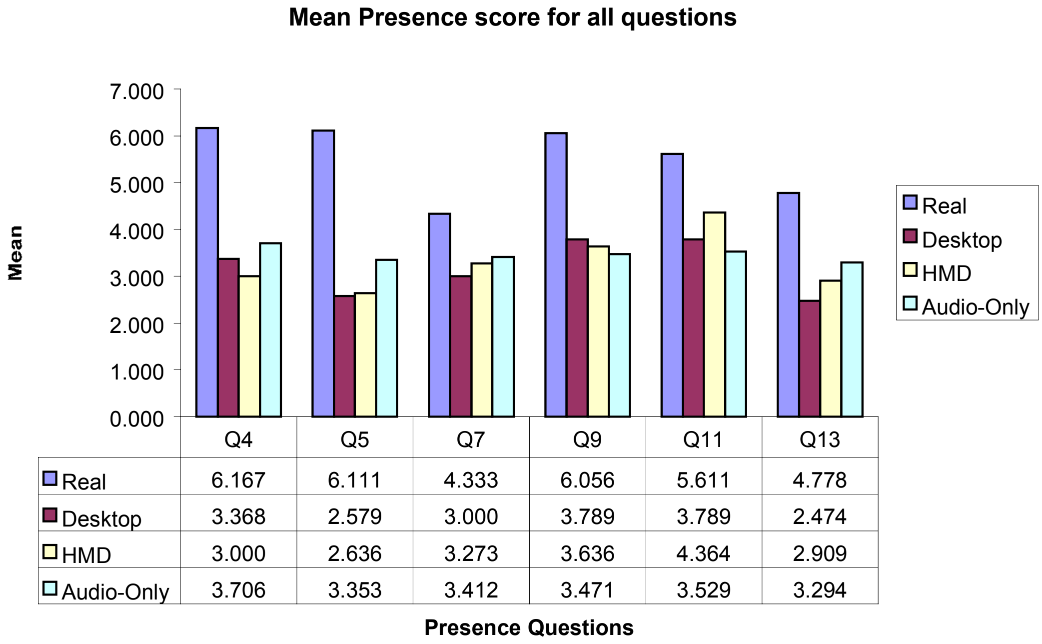
There was a significant overall effect of condition. Examining the results for each presence question separately (Appendix A1), only responses to one question did not prove to have a significant effect across conditions, $F(3,71)=1.857$, $p>0.05$. Replying to this particular question participants rated their impressions of the seminar room as images/sounds that they've seen/heard as opposed to somewhere they visited. Figure 4.12 shows the mean responses for the six questions assessing presence. These questions are included in Chapter 2. The full questionnaire can be found as used in this study in the Appendix A1. To calculate the presence scores, questions 4,5,7,9,11,13 out of this questionnaire were utilised. The overall effect for each of the six questions according to the ANOVA is as follows ($p<0.01$ indicates a stronger significance than $p<0.05$):

Question 4: $F(3,71)=15.407$, $p<0.01$
Question 5: $F(3,71)=19.971$, $p<0.01$
Question 7: $F(3,71)=1.857$, $p>0.05$
Question 9: $F(3,71)=12.282$, $p<0.01$

Question 11: $F(3,71)=6.875$, $p<0.01$
Question 13: $F(3,71)=6.093$, $p<0.01$

The level of presence overall was found to be significantly higher ($p<0.05$) for the real-world condition compared with the desktop, HMD and the audio-only condition, thereby supporting relevant predictions. However, there was no significant difference in presence between the desktop, HMD and audio-only conditions in any combination (Figure 4.13). Moreover, to the surprise of the author, there was a general tendency of the presence mean values for the audio-only condition to be slightly higher compared to the desktop and HMD condition. It could be assumed that since part of the task was not visual, the extracted audio from the real seminar which included a lot of 'realistic' ambient sound cues such as shifting of papers, coughing, etc. actually induced a strong 'sense' of presence. The computer graphics replica of the real seminar room was rendered flat-shaded. The unrealistic visual stimulus was not present to possibly emphasise the fact that this was a flat-shaded computer graphics simulation.

Valuable input towards explaining this phenomenon could be drawn from traditional theories of mental imagery as discussed in Chapter 3. Visual mental imagery is 'seeing' in the absence of the appropriate sensory input and it is different from perception that is the registration of 'physically' present stimuli. Hence, it could be argued that the audio-only condition provoked a high level of visual mental imagery. As mentioned in Chapter 3, it is proven that imagery plays an important role in memory and spatial and abstract reasoning as well as in skill learning. In the result section for the non-visual part of the task, no significant differences of accurate recall were revealed between the audio-only condition and either the desktop or HMD conditions, sustaining this theory. Gilkey & Weisenberger [GW95] in an investigation related to the sense of presence for the suddenly deafened adults and the implications for VEs, report that auditory cues are a crucial determinant of the sense of presence. They emphasise that the crucial element of auditory stimulation for creating a sense of presence may be the auditory background, comprising the incidental sounds made by objects in the environment rather than the communication that typically captures attention. The participants verified this as discussed in the qualitative analysis of the section 4.2.5.



	Q4	Q5	Q7	Q9	Q11	Q13
Real	6.167	6.111	4.333	6.056	5.611	4.778
Desktop	3.368	2.579	3.000	3.789	3.789	2.474
HMD	3.000	2.636	3.273	3.636	4.364	2.909
Audio-Only	3.706	3.353	3.412	3.471	3.529	3.294

Figure 4.12: Mean levels of reported Presence for all questions.

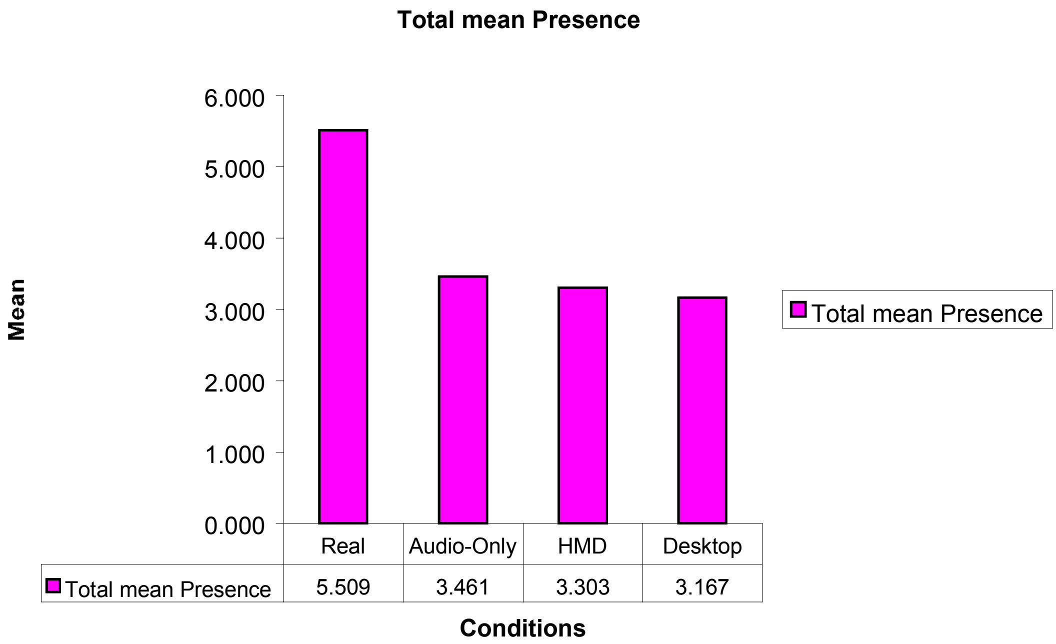


Figure 4.13: Total mean Presence scores across conditions.

There was not a significant difference of perceived presence between subjects experienced in playing computer games and the non-game players. Participants who preferred not to navigate significantly around the room had the highest level of presence, thus concurring with an observation by Snow & Williges [SW98] according to which perceived presence was not affected by the amount of possible interactions in the VE. The last question in the presence questionnaire requested the participants to rate their sense of losing track of time. It was observed that although presence was much higher in the real-world condition, the mean for this question was higher for the desktop condition followed by the real-world and the audio-only condition. The lowest mean rate was observed for the HMD condition and might be related to the level of discomfort of participants wearing the HMD. Therefore, assuming that losing track of time is linked with 'enjoyment' or 'involvement', the overall level of presence followed a different trend.

4.2.4 Aftereffects Results

The Head Mounted Display used for the HMD condition is a Hewlett Packard Laboratories working prototype (Figure 4.3, 4.4, 4.5). It is light, however, being a prototype it is not widely tested in the market. Since it allows for a generous peripheral vision, it is considered to be much safer than fully immersive VE HMDs which can give rise to severe vision problems and nausea. For the HMD condition, post exposure symptomatology reports were acquired and correlated with presence for any potential trends. The Simulator Sickness Questionnaire (SSQ) introduced by Kennedy et al. [KLBL93] was employed as a measure. A significant correlation was not revealed between the total SSQ score and the level of presence although there was a slight trend towards a significant negative correlation of –0.4 (Pearson's). A negative correlation suggests that if one variable tends to increase, the other decreases (co-varies). A negative correlation of –0.4, here, indicates that participants who experienced the highest level of presence experienced the lowest level of aftereffects but this result did not reveal a strong negative correlation. Generally, strength (-0.4, here) is a measure of the correlation but significance assesses how unlikely such a correlation was to occur under the null hypothesis (usually that the population correlation is zero) [Coo99]. The general levels of the SSQ symptoms for the HMD HP Labs prototype following a 15-minute exposure to a monocular, non-head tracked computer graphics world were generally quite low.

4.2.5 Qualitative Analysis of Participants Comments

The last part of the presence questionnaire (Appendix A1) included a blank sheet with the following question: 'Please write down any further comments that you wish to make about your experience. What things helped to give you a sense of 'really being' in the space, and what things acted to 'pull you out' of this?' Participants replied to this question directly, chose to report personal feelings and comments about their experience or preferred not to comment at all.

Participants reported the following aspects of the experiment that *enhanced* their feeling of 'being' in the experimental room as regards the desktop, HMD and audio-only conditions:

- Existence of background sounds such as people coughing and papers shifting: 20% of the participants in the desktop and HMD condition and 30% in the audio-only condition.
- Ability to navigate: 30% of the participants in the HMD condition.

- Photographic texture outside the window of the experimental room: 25% of the participants in the desktop condition.
- The general quality of the sound: 20% of the participants in the audio-only condition.
- Previous experience of being in any seminar room: 15% of the participants in the audio-only condition.

Participants also reported the following aspects of the experiment that *pulled them out* of a feeling of 'being' in the experimental room as regards the desktop, HMD and audio-only conditions:

- Photographic texture of lecturer not being animated: 65% of the participants in the desktop and HMD conditions.
- Need for more realistic rendering: 35% of participants in the desktop condition and 10% of participants in the HMD condition.
- HMD discomfort: 30% of the participants in the HMD condition.
- Not having a visual stimulus: 25% of the participants in the audio-only condition.
- Distracting surroundings - had to close eyes to have maximum effect: 30% of the participants in the audio-only condition.
- Will to have more freedom for navigation (for instance, zoom in and out): 10% of the participants in the desktop condition and 15% of the participants in the HMD condition.
- Lack of visualising other participants: 10% of the participants in the desktop condition.
- Lack of stereo imagery: 10% of the participants in the HMD condition.

Certain participants reported having visual images created by the realistic sound for the audio-only condition:

> 'I could visualise a man sitting in front of a class but looking back I can't remember colour, size, location. Like a dream, you know the image of what you visualised but you can't analyse it into segments and terms of the image'.

Other comments relevant to the desktop condition with an emphasis on sound were:

> 'Overall, I think the audio elements of the experiment had more impact on how 'real' the room was than the visual elements. While the ability to move around the room added reality, the fact that there is a feeling of not being in control of what was going on made it less real. If I could approach the lecturer

> and stood next to him in a real room, he would react but in the simulation he couldn't; this acted to make things less real'.

The novelty of the experience in terms of the computer graphics space was commented:

> 'I did find myself turning to look out of the window a couple of times but I think that was due to exploring something new (the 3D environment) and the 'new-ness' of that distracted me from the task at hand'.

4.2.6 Conclusions and Problems with Preliminary Study

In general, presence did not follow the same trend as task performance in all cases. For example, presence was significantly higher for the real-world condition compared to the desktop condition but that was not reflected on memory recall. There was no statistical difference between the scores for the spatial recall task between the real-world and the desktop condition. The presence questionnaire also revealed no significant difference between the technological conditions. This could mean that either these conditions do not have a varied effect on presence or that the measuring device, in this case, the presence questionnaire could not pick up that difference. Inherently, this could be an issue about the notion itself. There is an amount of ambiguity in terms of a scientific representation of the notion that might reflect onto any possible measuring instrument.

The incorporation of cognition-related measures, in this case, the report of the relevant memory awareness state for each item of the memory recall task offered a valuable input towards a more informative analysis. There was no statistical difference for the spatial memory task across conditions, but prior probabilities relevant to memory awareness states showed that the probability for an accurate response to fall under the 'remember' awareness state was higher for the HMD condition compared to the real-world condition. Since 'remember' responses are linked with visual mental imagery as a mechanism of retrieval, it could be argued that mental images and subsequent memory responses associated with the HMD condition are more 'vivid' or 'realistic' and that could have an effect on spatial perception retained in time. It is therefore suggested that usability studies involving only task performance measures while considering a possible design or technology such as the

Hewlett Packard HMD prototype, are not sufficient to form conclusions regarding the effectiveness of the design or hardware in question.

Although the preliminary study gave confidence in the memory semantics methodology based on the remember/know paradigm, the rendering used was basic and the spatial memory elements of the task were limited. The preliminary study demonstrated the potential of the memory semantics methodology as a simulation fidelity measure for VE applications in relation to the real world. This measure focuses on the cognitive processes participants employ in order to complete a memory task rather than on the actual scores of accurate completion commonly employed. The purpose of the preliminary study was to adjust the memory semantics methodology for VE immersive technology experimentation and reveal problems before a full study is made, therefore, the experimental design was not strict. The following problems arose during the preliminary study and needed to be addressed before the main study:

Complex task: The memory task employed in the preliminary study was too complex. The seminar-like experience resulted in an inadequate computer graphics representation. For instance, the lecturer who gave the seminar in the real world could not be simulated (or 'animated' accurately) but only included in the space as a static billboard. Gaze, facial expressions and gestures are significant elements of non-verbal communication [MC98]. Their absence in the computer graphics simulation introduces a confound in the experimental design with an unknown effect. The focus of the preliminary study was the spatial memory task. A large part of the memory task was not visual and although this resulted in interesting parallels between the visual and non-visual part of the task, a larger study should focus on a clearly defined spatial memory task. Additionally, in order to identify how mental processes influence recollection and retrieval over time, participants could be re-tested after a specified amount of time.

Rendering quality: As demonstrated in Chapter 2, most studies comparing real life situations with computer graphics simulations do not address the issue of the rendering quality, mostly because computational power is restricting the ability to provide better rendering. Although this is not necessarily a confound since compromises in quality are often necessary, the effect of the rendering quality level needs to be accounted for in any experimental design that strives to make comparisons with a real world situation. In the preliminary study, the rendering was kept flat-shaded, thus, achieving an impressive frame rate of 45 frames per second. The experimental room was not darkened and participants in all conditions utilising computer graphics imagery were aware of their surroundings. In the main study, it would be valuable to achieve photo-realistic rendering quality based on photometry measurements acquired in the real-world space towards accurate simulation of geometry and illumination for the computer graphics application by blocking real-world illumination.

Apparatus: In the preliminary study, there were slight differences in resolution and FoV between the desktop and HMD condition. The real world scene was perceived as a control condition and the FoV of participants was not restricted for this condition. In a more formal experimental design such differences should not exist. The HMD employed in this study was monocular and non-head tracked. In order to realise the full effect of virtual interfaces on spatial perception as revealed by the memory semantics methodology, more sophisticated VE technology should be employed such as stereo imagery and head tracking for navigation.

The preliminary study gave confidence in the methodology while establishing common procedures and conditions. While the results of the preliminary study look promising, action needs to be taken to remedy the shortcomings that have risen as explained above. Results from an improved framework can be analysed in more detail and confidence in order to draw conclusions that could identify cognitive variations of spatial perception derived by the memory semantics methodology.

4.3 Summary

In this chapter, the methodology based on the memory awareness states discussed in Chapter 3, for assessing the simulation fidelity of a computer graphics simulation of a real world situation is incorporated in a preliminary study. This methodology is focusing on the cognitive mental processes participants employ in order to complete a memory task in the real world, as opposed to using a computer graphics representation of the real world space displayed on a typical monitor as well as on a monocular HMD. The memory task includes a visual and a non-visual component. The goal of the preliminary study was to develop a robust framework based on the memory semantics methodology that could be subsequently applied to the main studies described in Chapter 5 that utilise a formal experimental design.

The level of reported presence was higher in the real world compared with the desktop, HMD and audio-only condition, however, there was no significant difference for presence between those conditions. A consistent positive correlation between presence and task performance was not revealed. Spatial memory recall and confidence socres did not prove to be significantly different across conditions. Memory awareness states' analysis, however, for the spatial task gave an invaluable insight into participants' strategies of retrieval, most interestingly across specific conditions where results for presence and accurate memory recall were not proven to be significantly different. In particular, statistically significant differences were revealed for the 'remember' awareness state linked with mental imagery and event-based recollections and for the 'familiar' awareness state linked with awareness of an event or an image that can not be accurately placed in time. One of the most interesting results showed a significantly higher proportion of correct 'remember' responses for the HMD condition compared with the real-world condition, revealing a high amount of visual mental imagery as a strategy of retrieval for this condition.

Building on the preliminary study, a complete experimental framework is presented in the next chapter. The result is a formal experimental design incorporating a spatial perception task with more elements and complexity. Also, sophisticated display technologies are incorporated such as stereo imagery and head tracking whilst allowing the investigation of the effects of proprioception on spatial memory.

Chapter 5

Main Experiments and Results

This chapter outlines the experimental methodology employed and the relevant results for the main studies in this thesis which compares spatial perception in a photorealistic VE, displayed on a desktop monitor as well as on a HMD in comparison to the real space it represents [MCTH2001], [Man2001a]. The preliminary study presented in Chapter 4 demonstrated the value and applicability of the memory awareness states methodology, however, it also highlighted the shortcomings of the informal experimental design. The actions taken to remedy those shortcomings are described in this chapter in the context of a larger study. The methodology based on the remember/know paradigm presented in Chapter 3 is employed in the testing process towards identifying the *mental processes* participants followed in order to complete a spatial memory task and it forms a framework for assessing the fidelity of a dynamic computer graphics representation.

The main study consists of two parts. Initially, the main spatial memory task study was conducted. The participants were required to complete the same task across conditions after a specified time of one week in order to identify the effect of the methodology and each condition on spatial awareness retained over time. The computer graphics rendering of the scene is near photorealistic (radiosity) incorporating photometry measurements acquired in the real space and controlled illumination. Experimental conditions include monocular and stereo computer graphics imagery displayed on a HMD. The effect of proprioception (mouse vs. head tracking) on spatial perception and cognitive states of recollection is explored.

5.1 Experimental Methodology

In this section, the experimental design is described in detail. Experimental design issues such as participants, apparatus and materials are analysed concentrating on extending the preliminary study's procedures and methodology. The design employs a larger number of elements of recollection and participants compared to the preliminary study and photorealistic rendering based on real-world photometric measurements. A visuo-spatial task was designed for the main experiments. Head-tracked HMD technology is incorporated displaying monocular and stereo imagery.

5.1.1 Methods

The experiments presented in this chapter employ a spatial memory task in order to investigate the degree to which a 'realistic' computer graphics simulation of a real world space could reveal similar recollection accuracy as well as cognitive awareness of this space as the real world space. Based on this process, a simulation fidelity tool is formed based on the actual performance of a spatial task but most importantly, on the mental processes participants employ towards completing the task. A spatial task of orientation and object recognition was adopted. The memory semantics methodology described in Chapter 3 and applied in the preliminary study of Chapter 4 was incorporated in the experimental design. It is challenging to determine the degree to which the cognitive strategies of retrieval associated with spatial recollections remain unchanged in a VE simulation in comparison to the real world scene being represented.

Five groups of 21 participants were recruited to participate in the main study, from the University of Bristol, UK undergraduate and M.Sc. student population and they received course credits for their participation. None of these participants had taken part in the preliminary study of Chapter 4. 80% of the subjects from each group were male. All use computers a great deal in their daily activities. Participants were randomly assigned to each group. A between-subject design was utilised balancing groups for age and gender. Participants in all conditions were informed that they could withdraw from participation at any time during the experiments and they were naive as to the purpose of the experiment. Participants had either normal or corrected-to-normal vision. According to the group they were assigned to, participants completed the same spatial memory task in one of the following conditions:

1) In reality, wearing custom made goggles to restrict their FoV, allowing for monocular vision; referred to as the **real-world condition**.
2) Using a photorealistic computer graphics simulation on a monocular head-tracked HMD; referred to as the **HMD mono head tracked condition**.
3) Using the same application on a stereo head-tracked HMD; referred to as the **HMD stereo head tracked condition**.
4) Using the same application on a monocular HMD with a mouse interface; referred to as the **HMD mono mouse condition**.
5) Using the same application displayed on a typical monocular desktop monitor with a mouse interface, wearing the same restrictive goggles as in the real-world condition; referred to as the **desktop condition**.

The participants completed a spatial memory task with a confidence and awareness states report, a presence questionnaire, an SSQ questionnaire before and after exposure and a questionnaire on subjective responses to lighting as discussed in Chapter 2. A week after their experience, all participants were re-tested on the same spatial memory task across conditions.

Hypotheses / Prediction

Accurate memory recall was expected to be similar across conditions due to the high quality of the computer graphics rendering. The scope of the main studies was to identify variations on the cognitive strategies of recollection for the main task as well as the retest and associations between the awareness states distributions. The emphasis was on replicating relevant results in the preliminary study which revealed that the proportion of correct 'remember' responses was higher for the HMD mono mouse condition, in this case including varied HMD interfaces for the technological conditions. The level of presence was expected to be significantly higher in the real-world condition compared to the technological conditions and significant positive correlations of the presence dataset with the subjective responses to lighting dataset were predicted.

5.1.1.1 The Real Environment

The real environment consisted of a four by four meters room (Figure 5.1). Each wall of this room had a different landmark; one wall consisted of a door and shelves, one wall of a door and a greenboard, the third wall of a whiteboard and the fourth of smaller shelves on both its ends. The existing window in the room was firmly covered with black lining to keep natural light out in order to control the illumination in the experimental space. The light fixtures in the room were replaced with a standard incandescent bulb (assumed diffuse, light emission in all directions). Several tables were placed close to the walls and 21 primitive objects of approximately the same size (seven boxes, seven spheres and seven pyramids) were scattered around the room, on the tables and shelves. All the objects were painted one shade of blue using the same diffuse paint. A swivel chair was placed in the middle of the room.

5.1.1.2 The Computer Graphics Simulation

The computer graphics representation of the real environment was created using the 3D Studio MAX modelling suite [3ds2001] and Lightscape radiosity software [Li2001]. To describe a physical environment, the illumination and interaction of the light source with the environment needs to be included as well. A lighting simulation program takes as input the geometry of the environment, the properties of the light sources and the material characteristics of the surfaces and objects. The geometry in the real room was measured using a regular tape measure with accuracy of the order of one centimetre.

Figure 5.1: The real world room; the experimental space for the real-world condition.

A photometry instrument (Minolta Spot Chroma meter CS-100) was employed to measure the chromaticity CIE(x,y) and luminance (Y) values of the light and materials in the real room. Luminance is a photometric term as explained in Chapter 2; its units are candelas per square metre (cd/m^2). It relates to the quality of a colour that most resembles the human's notion of brightness. Bright colours are generally of a high luminance and dark colours are generally of a low luminance. The Minolta chroma meter is a compact, tristimulus colorimeter for non-contact measurements of light sources or reflective surfaces. The illuminant (light source) was measured by placing a white sheet of paper in a specific position. Most of the materials (walls, objects, shelves, floor, plugframes) were measured at the same position. To ensure accuracy five measurements were recorded for each material, the highest and lowest luminance magnitudes were discarded and an average was calculated of the remaining three triplets. The Lightscape radiosity rendering system uses RGB tristimulus values to describe surface characteristics. The values obtained for the illuminant and surfaces in the scene with the chromameter were converted from luminance and chromaticity co-ordinates to tristimulus RGB values. Measured chromaticity values were converted to RGB triplets by applying a matrix based on the chromaticity co-ordinates of the monitor phosphors, taking into account the illuminant measurements [Tra91]. Figure 5.2 shows the 3D model of the room and the outlined polygons before the radiosity rendering. Figure 5.3 shows the final radiosity solution of the model; the polygon count has obviously risen, especially in areas requiring a large amount of shadow computation.

The usefulness of the CIE(x,y) representation is that it allows colour specification in one language, however, equal geometric steps of CIE(x,y) space do not correspond to equal perceptual steps. For instance, humans are more sensitive to steps in the yellow region of the colour distribution than to steps in the blue region. Before specifying display colours in CIE(x,y) space, it is necessary to compute the tristimulus matrix of the display in question. In order to compute the RGB tristimulus matrix, the chromaticity co-ordinates of the three display phosphors in CIE(x,y) space is required. In addition, the chromaticity co-ordinates of the white that the three phosphors of the display produce when turned on at their maximum are also required.

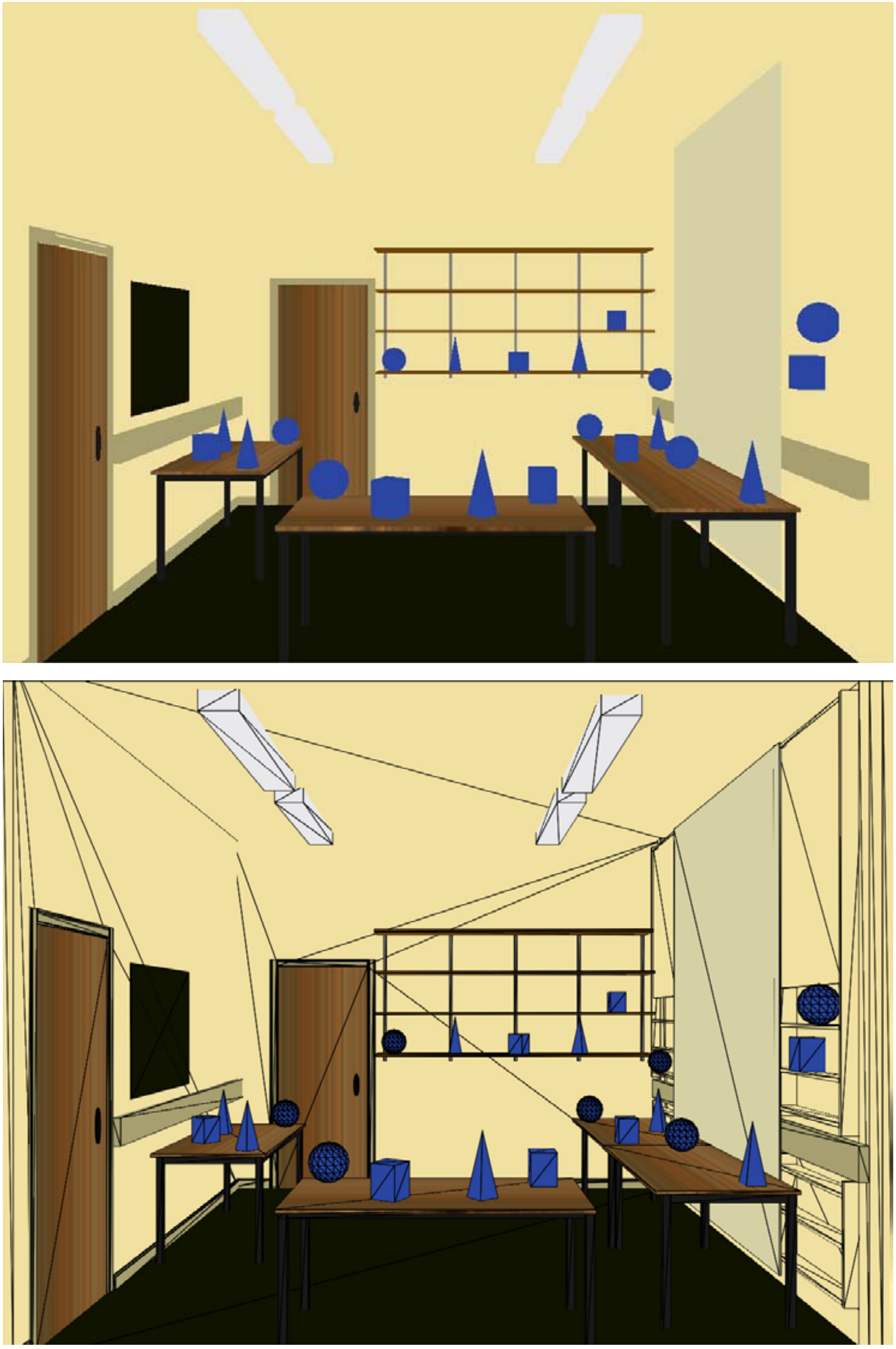


Figure 5.2: 3D computer graphics model of the real-world room, flat shaded and geometric mesh outline.

Figure 5.3: The radiosity rendering and mesh outline.

Generally, the RGB system is a means for describing colours on a display monitor. It is not a means for describing light energy in an environment. The RGB system does not take into account the energy that is produced in the physical world in terms of the distribution over wavelength and also how the Human Visual System (HVS) responds to this distribution.

For the final measurements, the illuminant had to be taken into account since measuring a diffuse surface under a given light source results in Yxy values which include the contribution of the light source itself. The CIE (1931) colour space is based on colour matching functions derived by human experimentation and it incorporated the trichromacy of HVS. Incandescent bulbs are quite orange and fluorescent light is quite green, however, the HVS perceives light in relative values and not as absolute measurements such as the ones out of the chromameter. For example, if 1000 is the luminance in the real world, 100 the luminance of a real-world material and 100 the luminance in the computer graphics simulation, then the luminance for the simulated material needs to be 10 for the same *ratio* to be preserved. The colour constancy attribute of the HVS, generally, corrects for these effects and is responsible for humans perceiving a white sheet of paper as white under a wide range of illumination (Chapter 2). If a participant is *immersed* into a synthetic space on a display, theoretically, this should be true as well, however, the small size of the displays prevent colour constancy from happening. In relevant calculations for simulating real-world illumination in a synthetic world, therefore, colour constancy needs to be corrected in the rendering process since the HVS does not function as in the real world due to the nature of the displays.

Generally, all the above principles are quite complex issues related to colour vision and how the brain deals with perceptual constancies as discussed in Chapter 2 and are not fully understood. In this study, the illuminant in the real room as measured with a white sheet of paper was taken into account in the conversions of the CIE(x,y) co-ordinates to RGB for all the materials measured in the real experimental room. The colour of the illuminant in RGB values was set as (1,1,1) for the radiosity rendering, e.g. white. The relevant calculations are as follows:

Output of the chromameter: x,y chromaticity co-ordinates and Y luminance.

Using measured x,y values, the z for each triplet can be calculated using the relationship:

$$z = 1 - x - y$$

The tristimulus matrix of the display has the following form, C^{-1} is the inverse of this matrix (r is red, g is green, b is blue):

$$[C] = \begin{vmatrix} x_r & x_g & x_b \\ y_r & y_g & y_b \\ z_r & z_g & z_b \end{vmatrix}$$

(x_w, y_w, z_w) are the CIE chromaticity co-ordinates of the white on the display. The tristimulus values X_w, Y_w, Z_w need to be calculated:

$$X_w = \frac{x_w}{y_w}, \ Y_w = 1, \ Z_w = \frac{z_w}{y_w}$$

$$\begin{vmatrix} V_1 \\ V_2 \\ V_3 \end{vmatrix} = [\,C^{-1}\,] \begin{vmatrix} X_w \\ Y_w \\ Z_w \end{vmatrix}$$

$$[T] = [C] \begin{vmatrix} V_1 & 0 & 0 \\ 0 & V_2 & 0 \\ 0 & 0 & V_3 \end{vmatrix}$$, where T^{-1} is the inverse of the T matrix calculated.

To compute the XYZ values given co-ordinates x,y,Y:

$$X = x\,\frac{Y}{y}$$

$$Y = Y$$

$$Z = (1 - x - y)\,\frac{Y}{y}$$

Then finally:

$$\begin{vmatrix} R \\ G \\ B \end{vmatrix} = [\,\mathrm{T}^{-1}\,] \begin{vmatrix} X \\ Y \\ Z \end{vmatrix}$$

In order to render the scene, the materials' diffuse colour needs to be specified not the colour observed under a particular light source. The colour is the amount of light at each wavelength that is not absorbed. The final colour for each measured material in the scene is estimated by dividing its RGB value by the RGB value of the observed white in the scene, which is the colour of the light source in the scene.

The material measurements in CIE(x,y) using the Minolta chromameter were quite accurate for the small samples on which they were made. However, as this is a room in daily use some variations exist in all of the surfaces due to texture, ageing and dirt. Using the relevant geometry, surfaces and illuminant measurements converted to RGB triplets as input, the rendered model was created using the Lightscape radiosity rendering system. Lightscape uses RGB tristimulus values to describe surface characteristics. Table 5.1 shows the relevant conversions of the CIE (x,y) co-ordinates measured in the real room to RGB values used in the radiosity rendering process (Figure 5.3). The final radiosity solution consisted of a finely meshed model which could be interactively manipulated. This was the basis for the application displayed on the desktop monitor and on the HMD. The application was implemented using the Sense8 WorldUp system and API [Wor2001].

The viewpoint for navigation was set up in the middle of the room and interaction was limited to a 360 degrees circle around that viewpoint horizontally and approximately 180 degrees vertically in order to simulate participants' head and eye movement on the swivel chair in the real room. Participants in the desktop condition and in the HMD mono mouse condition used a common mouse for navigation with the restrictions described above, thus, were sitting still on a chair. The participants in the HMD mono and stereo head tracked conditions actually utilised head tracking for navigation. Their positioning on the swivel chair and movements resembled the equivalent navigation for participants in the real-world condition (360 degrees rotation). The frame rate was retained at 14 frames per second across all conditions.

Although this is not a particularly high frame rate, it was considered adequate. The same frame rate was retained across conditions including the HMD stereo head tracked condition. Texture mapping was applied only on the doors and tables in the room. The resolution was 640*480 (HMD maximum resolution) and the FoV was 30 degrees across all conditions including the real-world one with the goggles fitted.

Surface	**Y**	**x**	**y**		**R**	**G**	**B**
Illuminant	23.4	0.470	0.396		1	0.34	0.10
Greenboard	2.13	0.412	0.487		0.0468	0.1310	0.0258
Walls	20.83	0.494	0.407		0.9461	0.8540	0.4888
Plug-frame	14.9	0.484	0.405		0.6541	0.6276	0.4410
Objects	5.4	0.381	0.362		0.1638	0.2762	0.6354
Floor	1.26	0.497	0.464		0.0475	0.0612	-0.0136
Whiteboard	20.7	0.482	0.405		0.9012	0.8778	0.6326
Shelves	8.17	0.534	0.402		0.4385	0.2808	0.0606
Lining	19.4	0.479	0.403		0.8400	0.8250	0.6476

Table 5.1: Conversions of CIE(x,y) co-ordinates measured in the real room to RGB values used in the radiosity renderer.

5.1.1.3 The Displays

The desktop monitor and the HMD were gamma corrected using the Minolta Spot Chromameter CS-100 in order to acquire relevant luminance readings. When accurate colour specification is required as is often the case in scientific applications, the non-linear relationship between display luminance and voltage is a significant source of error. In an ideal display the function relating display luminance and voltage would be linear. Mathematically, if D is the value sent to the digital-to-analog convertor and L is the luminance then this ideal relationship could be expressed as:

$$L = mD + c$$

where m, c are constants.

The ideal relationship simplifies to:

$$L = D$$

In practice though, the function is found to be non-linear. The actual luminance is typically modelled by a function of the form:

$$L = D^{\gamma}$$

where γ is a constant that depends, among other things, on the luminance and contrast settings of the screen. This is often described as a *gamma function*. *Gamma correction* is the process where voltage values are computed to generate required luminances. The transformations described above from CIE(x,y) values yield three RGB values; those RGB values represent proportions of the maximum luminance. For instance, RGB(0.5 0.5 0.5) means that each primary colour should be set at half of each maximum luminance. In order not to violate the linearity assumption, it is necessary to gamma correct or linearise the system. Otherwise the specification RGB(0.5 0.5 0.5) may set each primary at only about 20% of each maximum luminance. This is theoretically achieved by inverting the function by computing:

$$D_c = D^{1/\gamma}$$

where D_c is the gamma corrected value that should be passed to the digital to analog convertor, D is the value before gamma correction and γ is a constant.

For a display that is not gamma corrected, a gamma may be assumed between 2.0 and 3.5. There are crude ways of measuring it roughly by matching grey patches on the screen from a specific distance; these represent 0.2 intervals of gamma values for a specific stimulus [Tra91]. But where accurate colour specification is required as in this study, the function should be measured precisely with a luminance meter and, therefore, gamma should be estimated formally. The graphics card drivers utilised in the set-up for the study presented in this chapter incorporated gamma settings so gamma correction could be achieved through hardware. The gamma input for the graphics card was manipulated so as to achieve a linear relationship between voltage and luminance. A gamma of 1.1 was achieved for the desktop monitor and HMD and was confirmed by flooding the display with grey patches starting from lighter towards darker in intervals of 20 units of luminance. In a dark room, the relevant luminances were measured with the chromameter confirming this linear relationship. The monitor and the HMD were both kept turned on after the gamma correction process. Theoretically, every time a display is switched off its gamma function slightly changes and needs to be recalculated; it takes 2 hours to 'warm up' to its final value.

5.1.2 Materials

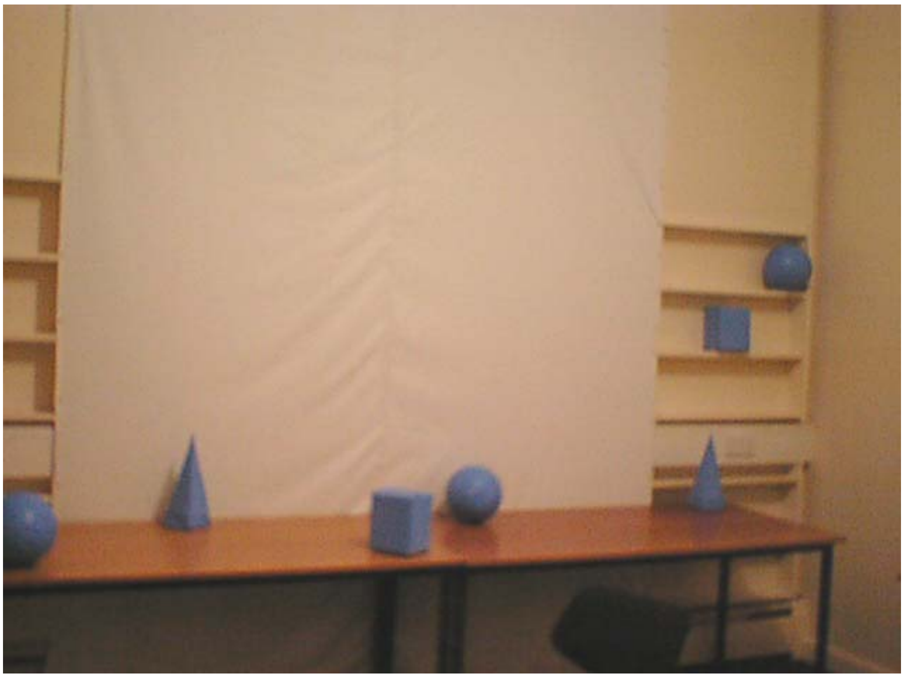
The five groups of participants were asked to fill in the same set of questionnaires. This set included the SSQ questionnaire before and after the task in order to detect any possible variations on symptomatology in relation to the participants' initial state, the spatial memory task and memory awareness states questionnaire, the presence questionnaire and the subjective responses to lighting questionnaire. Participants completed the spatial memory task for a second time, a week after the initial study. It was important to establish whether they retained spatial information in time as well as to explore any variations in their cognitive awareness across conditions related to the initial study. For a complete pack of all the materials involved in this study, please see Appendix A2.

Memory recall task

The questionnaire relevant to the task for the real, desktop and HMD conditions was designed to test the participants' memory recall of the positions of the 21 objects in the room. A diagram for each wall in the room, stylistically drawn using a professional architectural design package, i.e. Autocad [Au2001] included numbered positions of objects in various locations. The diagrams were administered together with the task questionnaire which consisted of 21 multiple-choice questions for the 21 objects in the scene, the same across all conditions (Figures 5.4, 5.5, 5.6, 5.7). Every question had three possible answers (box, sphere, and pyramid) and it included a confidence measure with five possible states: No confidence, Low confidence, Moderate confidence, Confident, Certain. It also included an awareness measure based on the memory semantics methodology with four choices: Remember, Know, Familiar and Guess. The participants had to orient themselves according to the diagrams of the walls provided and their recollection of the room and subsequently report on the shape of the object in each numbered position on the diagram, starting with the positions they were more confident they remembered. The experimental design, thus, of the actual task questionnaire did not force participants to start from a specified position in the room offering the capability to report, initially, their more confident recollections. Prior to the main study a pilot study was conducted in order to determine the number of objects and, therefore, the number of questions of recall as well as the actual amount of time that participants would be allowed to view the room across all conditions. It was significant to avoid possible floor or ceiling effects (the task being too easy or too hard) and that was a function of the viewing time as well as the number of objects that had to be recalled. It was determined that by including 21 objects in the room for 3 minutes viewing time participants in the pilot study could recall an average of 11 objects accurately for the 3 minutes specified, which was above chance, thus, this was a satisfactory combination.

Prior to filling out the core of the task questionnaire, participants were given instructions that were designed to explain what each of the memory awareness states depicted. The instructions in this study were modified to be shorter that the ones employed in the preliminary study as follows:

- REMEMBER means that you can visualise clearly the object in the room in your head, in that particular location. You virtually ‘see’ again elements of the room in your mind.
- KNOW means that you just ‘know’ the correct answer and the alternative you have selected just ‘stood out’ from the choices available. In this case you can’t visualise the specific image or information in your mind.
- FAMILIAR means that you did not remember a specific instance, nor do you know the answer. It may seem or feel more familiar than any of the other alternatives.
- GUESS means that you may not have remembered, known, or felt that the choice you selected have been familiar. You may have made a guess, possibly an informed guess, e.g. you have selected the one that looks least unlikely.



Start Position

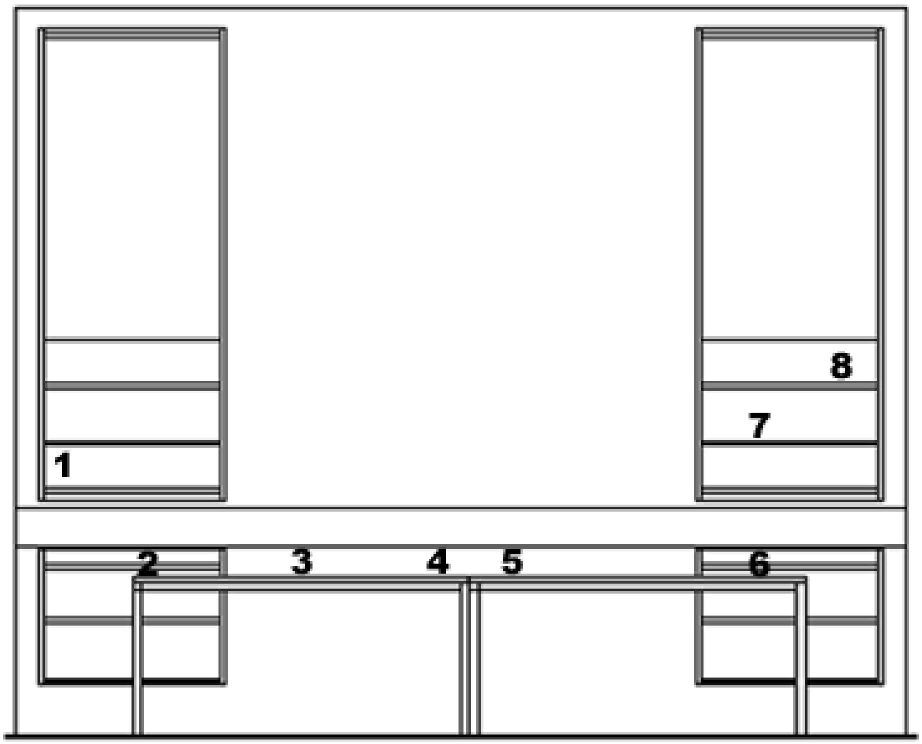


Figure 5.4: Wall in the real-world room representing start position of viewing and relevant diagram.

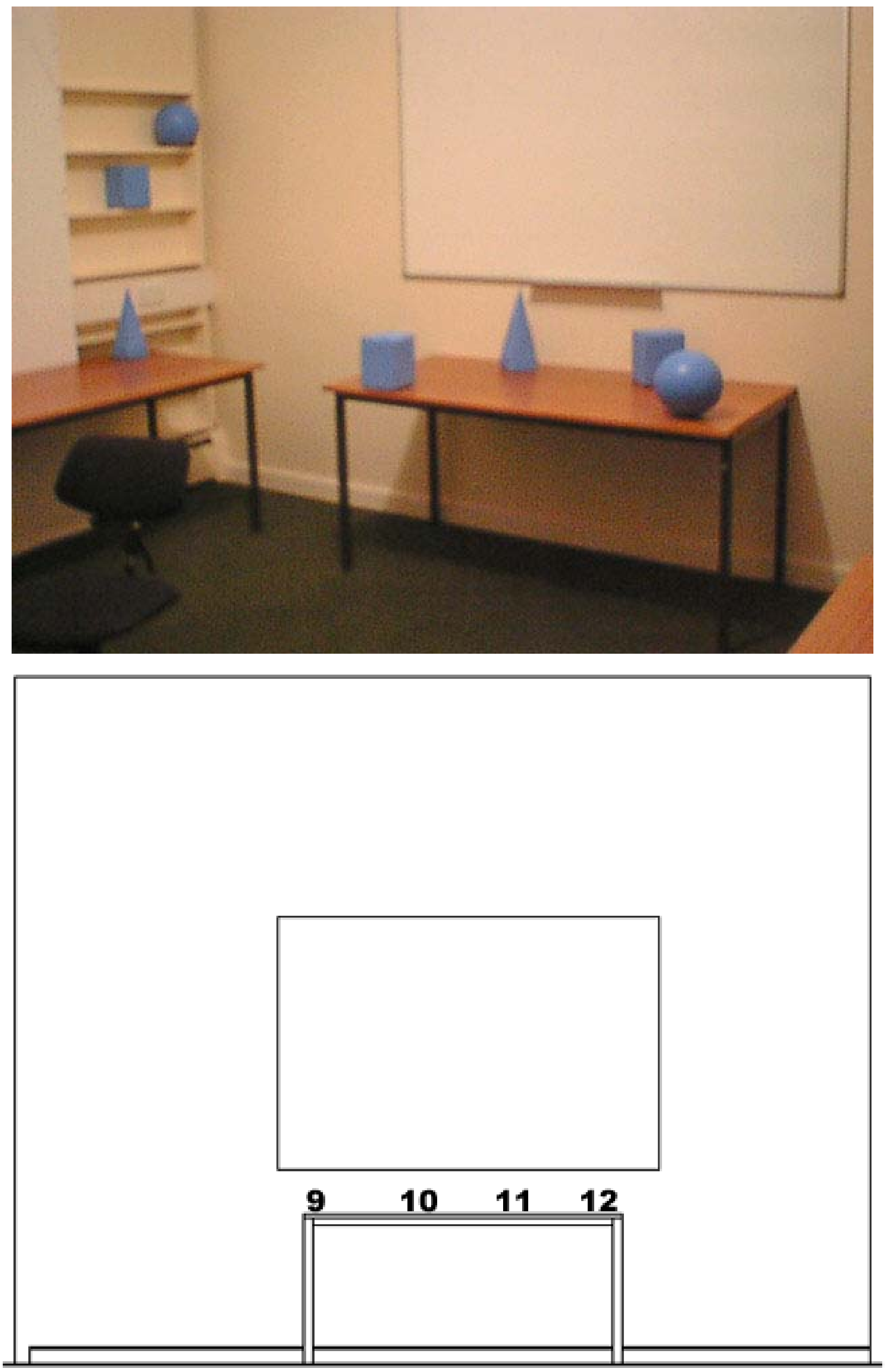


Figure 5.5: 2^{nd} wall in the real-world room and relevant diagram.

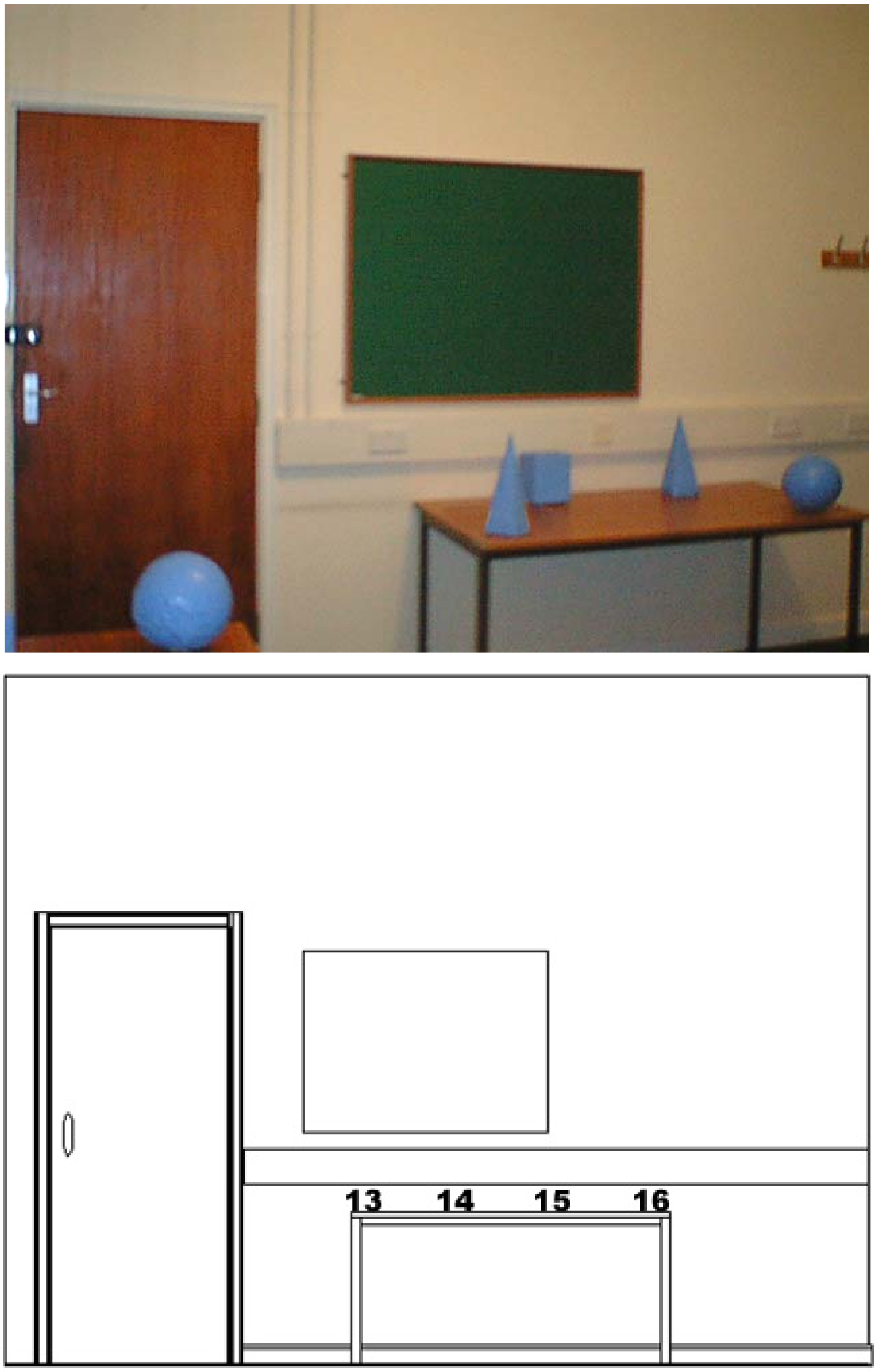


Figure 5.6: 3rd wall in the real-world room and relevant diagram.

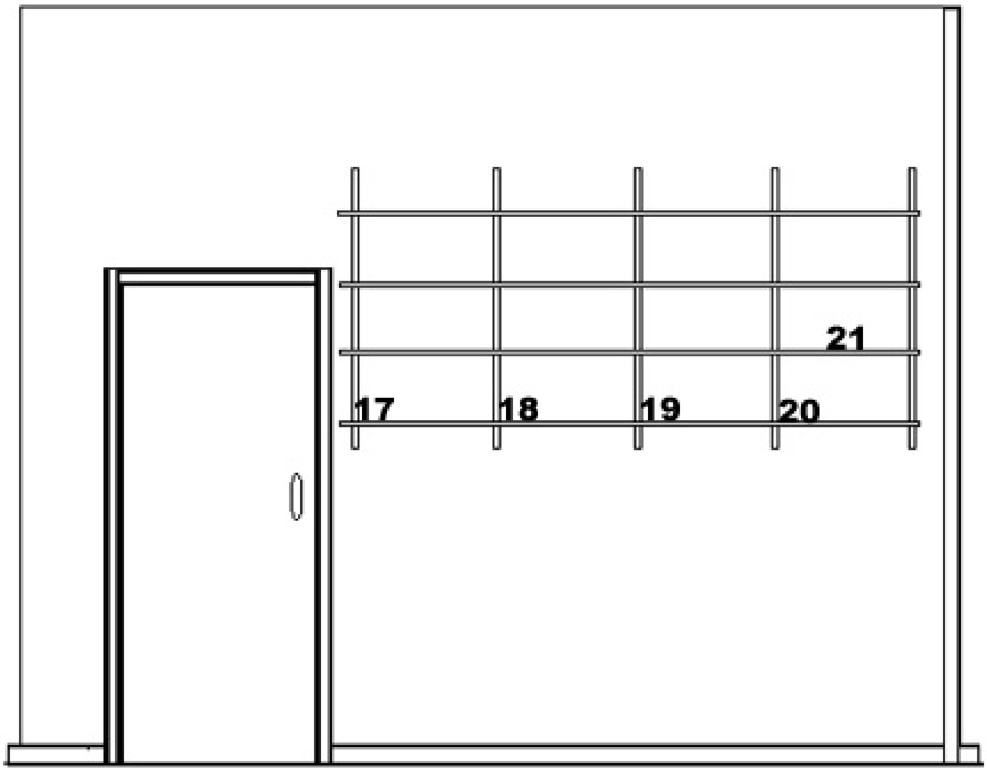


Figure 5.7: 4th wall in the real-world room and relevant diagram.

Presence

The second questionnaire was designed to measure the level of presence on a Likert 7-point scale. The questions used in the Slater et al. study [SSMM98] were adopted in order to assess the level of presence of the participants in each condition as in the preliminary study described in Chapter 4.

Subjective responses to lighting

The subjective responses to lighting questionnaire was administered in order to correlate its results with the results of the presence questionnaire [Rus2000]. Lighting designs may intentionally or unintentionally function more actively as selective intervention in human visual experiences: guiding circulation, focusing attention and otherwise affecting impressions of a room or activity. A relevant theory argues that light cues signal subjective associations or impressions and that the direction of these impressions is somewhat independent of the room in which the light cues are viewed [Fly77]. Similarly to this theory, it could be hypothesised that by accurately simulating the illumination in the real world to match the illumination in a synthetic space across displays such as desktop monitors and HMDs, subjective responses to lighting will not vary across conditions if the illumination is simulated accurately. At the same time, one wonders if related subjective reports are only a function of the accuracy of the light simulation from real to synthetic spaces and how perception of synthetic spaces, in that sense, differ to their real counterparts. This investigation was an addition to the main study. Although this exploration is not central in the experimental design, a subjective responses to lighting questionnaire is included in the testing process [Rus2000]. The rendering is retained the same across all displays involved in this study. Testing participants self-report of subjective impressions for the same scene across conditions could provide a means of validation for this method. Theoretically, if the participants' response is similar across conditions for the same dynamic computer graphics scene this could be a step towards validating the metric which could be subsequently used for assessing subjective responses to varied lighting or rendering quality scenes.

SSQ questionnaire

The widely used Simulator Sickness questionnaire (SSQ) was administered before and following participants' exposure to the VE for the HMD condition only. 16

symptoms were employed indicated in the Kennedy et al. study [KLBL93]. The questionnaire design is based on three components: Nausea, Oculomotor problems and Disorientation. Participants report the degree to which they experience each of the above symptoms as one of ‘none’, ‘slight’, ‘moderate’ and ‘severe’. These are scored respectively as 0,1,2,3 and analysed as discussed in Chapter 2.

5.1.3 Procedures

The real-world condition (Figure 5.8)

The SSQ questionnaire was initially administered to each of the participants of the group who completed the task in reality in order to test for simulator sickness symptoms before the experiment took place. Following this procedure they were asked to wear any glasses or contact lenses they normally use when they have to focus at 2 meters distance. The experimental process relied on self-report for this judgement due to the nature of the task. Subsequently, their dominant eye was identified. In order to make that judgement, participants were asked to align their index fingers on a straight line with both their eyes open and then without moving them, close their left and right eye and report with which eye they perceived their index fingers better aligned. This is called the ‘sighting’ test and it is widely used. The same viewing position for all the participants was set by manipulating the height of the swivel chair according to the individual. They wore appropriate goggles that restricted their FoV to 30 degrees to match the desktop and HMD FoV allowing for monocular vision through only the dominant eye (Figure 5.9). Participants were clearly instructed that they would be guided to a room where they will spend three minutes observing by rotating the swivel chair they will sit on, placed in the middle of the room. The participants were advised to take a careful look at the space. Each one of them was then guided into the experimental room and sat on a swivel chair in the middle of the room. The chair carried an electronic device, incorporating a digital compass which would monitor participants’ rotation in the specified time of three minutes, storing two angle measurements each second (Figure 5.10). Participants were, therefore, monitored in terms of their navigation and idle time while observing the room. The scope of this strategy was to correlate these results with the actual memory recall and awareness results as well as with the relevant navigation patterns of participants using the head tracked HMD or using the mouse. This dataset could also identify differences between navigation with proprioception cues (rotating the actual chair) as opposed to mouse navigation (rotating with the mouse). After the allowed specified time of three minutes, the participant was guided to the test room where the questionnaire pack was administered together with the appropriate instructions.

The technological conditions (Figures 5.11, 5.12)

The computer graphics application was displayed on a Kaiser Pro-View 30, gamma corrected HMD. The viewpoint was set in the middle of the room and navigation was restricted to a 360 degrees circle around that viewpoint. The geometric FoV was calculated to be the same as the visual angle, through the goggles, in the real room. For the HMD monocular conditions (head-tracked and non-head-tracked) the dominant eye was identified and the appropriate screen of the HMD was covered allowing for vision only through the dominant eye. For the HMD stereo head tracked condition, each participant's interpupilary distance (IPD) was measured and the stereo application's parallax was set accordingly. For the head tracked conditions (mono and stereo) similar measurements of rotation by means of the electronic device on the swivel chair were monitored. For the HMD mono mouse condition horizontal rotation was monitored through software according to mouse movements.

For the desktop condition, the dominant eye for each individual was identified, the appropriate goggles were subsequently worn as in the real-world condition and each participant used the application displayed on a gamma corrected, typical 21-inch desktop monitor. The frame of the monitor was covered with black cardboard to achieve a foreground occlusion effect resulting in a strong sense of depth as discussed in Chapter 2. Similarly for the desktop condition, horizontal rotation was monitored through software according to mouse movements.

There was no other source of light besides the HMD or desktop display during experimentation across all four technological conditions. The initial photometry measurements of the materials in the real room and relevant calculations for their transfer in the radiosity rendering engine took advantage of the colour and lightness constancy ability of the human visual system as analysed in the previous section.

Figure 5.8: Real-world condition.

Figure 5.9: Restrictive goggles allowing for a 30 degrees FoV.

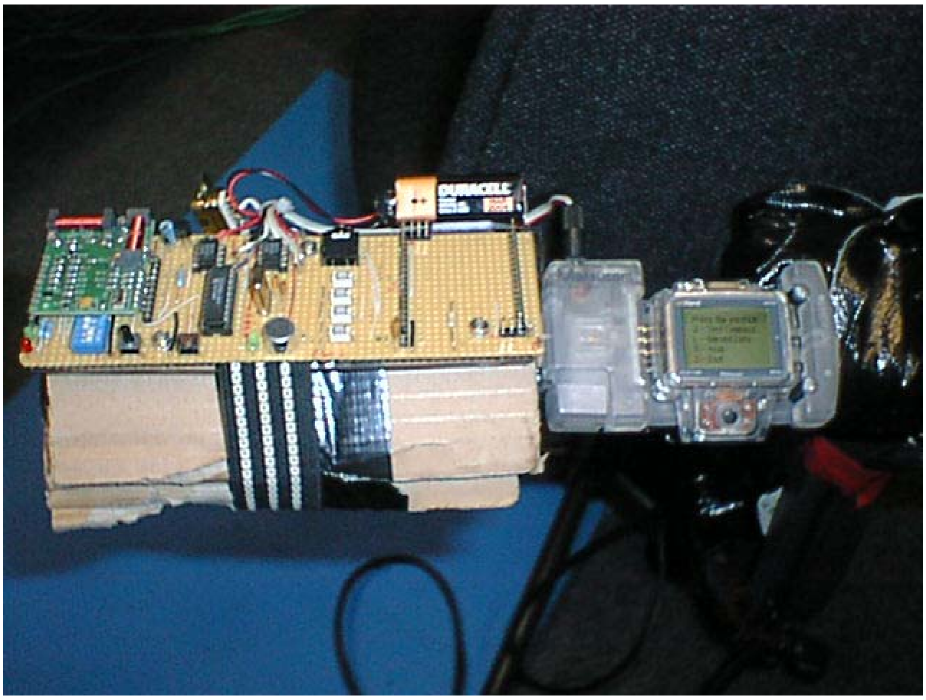


Figure 5.10: Digital compass on the chair monitoring navigation/idle time of movement.

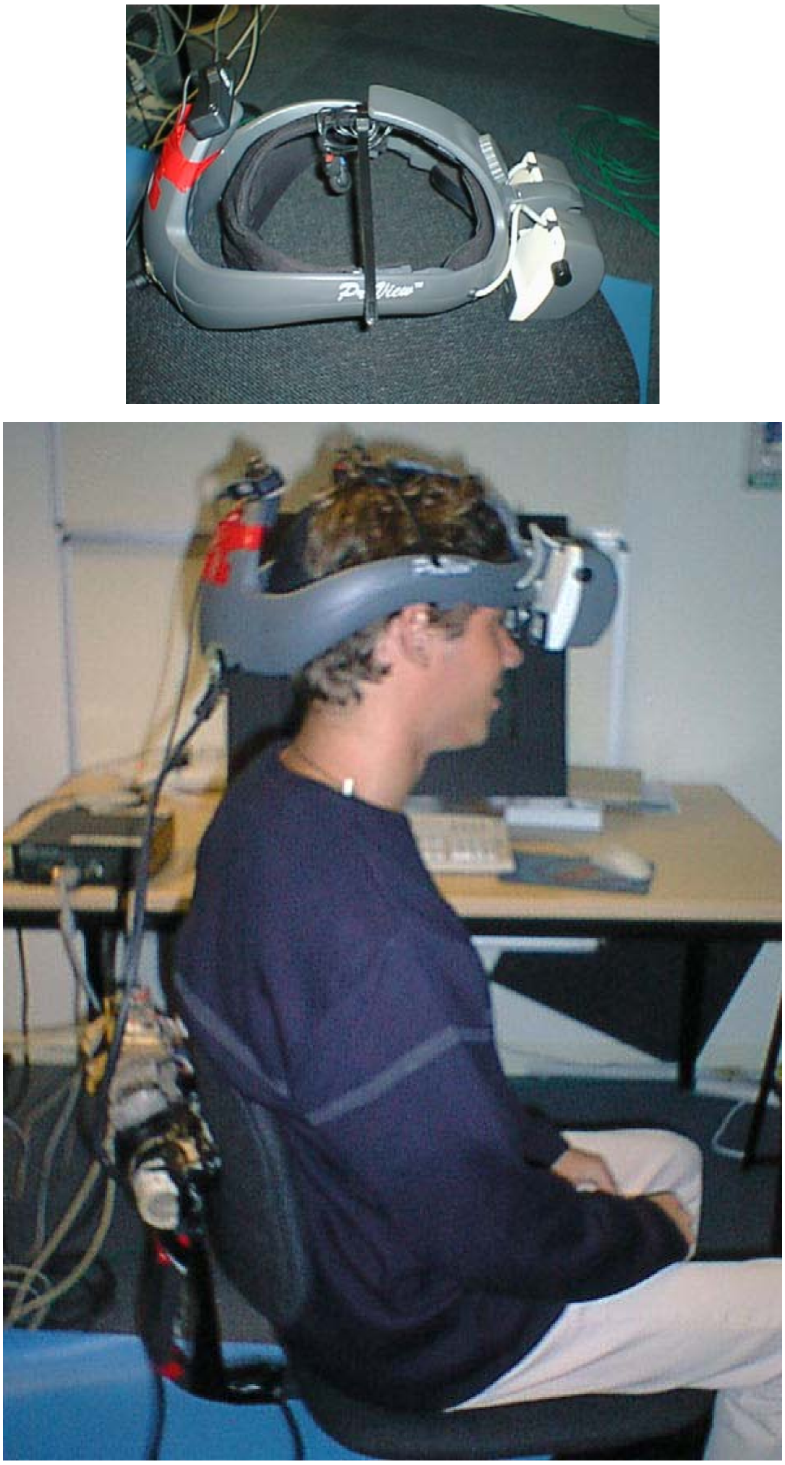


Figure 5.11: The HMD conditions.

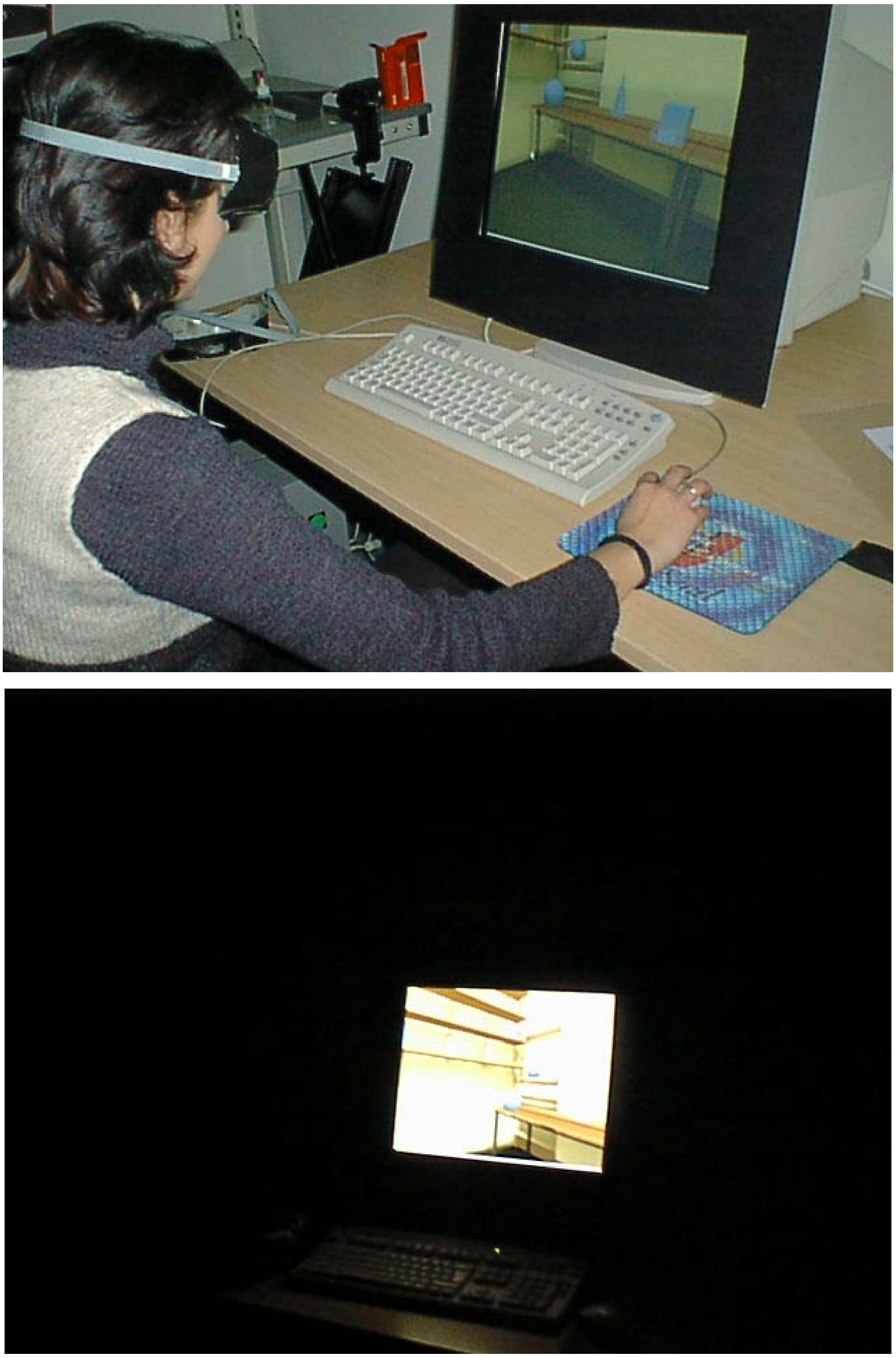


Figure 5.12: The desktop condition.

It has to be noted, though, that although the experimental space was dark there was some light shed from the HMD or the desktop display, illuminating the actual surroundings of the experimental space. Certain participants noted this effect as discussed in the qualitative analysis of section 5.2.5. Generally, the instructions before and after the completion of the experiment were similar across all conditions. A pilot study prior to the main study was concerned with the experimental testing process and relevant instructions before and after exposure, fine-tuning the procedures involved. To summarise, this is the experimental script for the experimental process:

- Administer Name/Time sheet.
- Ask if they have corrected to normal vision. Tell them to wear any optical correction they would normally wear if they look 2 metres away.
- Administer SSQ.
- Identification of dominant eye.
- Height investigation.
- Interpupilary distance investigation for stereo condition.
- Check if they are carrying electronic devices and large amount of metal/keys to avoid interference with the digital compass.
- Ask them if they know anything about this experiment and if they discussed it with anybody.
- Instructions: 'I am going to leave you for three minutes to view a room. You are not allowed to walk around in the room; you can only rotate the swivel chair you are sitting on (or you can use the mouse to rotate around the specified viewpoint and also look up and down). I would like you to look around as much as you can for this time. Please, be observant of the layout of the room and the symmetry and positioning of all the objects you are going to see scattered around. When the specified time has passed you are going to be asked to complete certain questionnaires in a different room on positioning and symmetries of the different object categories you are going to see. The entire process should take no more than twenty-five minutes. So, please, be observant of the layout of the room and of all the objects in it'.
- Wearing the appropriate goggles, with the slot shut.

< *The participant is guided in the experimental room and the digital compass/tracker where appropriate, is initialised. The participant is advised to start looking and vision is allowed by removing the cardboard preventing viewing on the goggles or HMD. After three minutes, the participant is guided in the test room.* >

5.2 Results and Discussion

5.2.1 Spatial Memory Recall and Memory Awareness States Results

The participants completed the same spatial memory task including confidence and awareness reports across the five conditions. The memory recall scores as well as the confidence scores were analysed using **AN**alysis of **VA**riance (ANOVA). ANOVA is a powerful set of procedures used for testing significance where two or more conditions are used. Significance decisions involve rejecting or retaining the null hypothesis (which claims that groups are identical). The null hypothesis is rejected when the probability that a result occurring under it is less than .05. Generally, ANOVA compares the variance of the sample means (between groups variance) with the within groups variance. Within groups variance is calculated by taking the average of the variances within each sample around its mean. As in the preliminary study, a quantity and quality analysis was conducted regarding the memory awareness states statistical investigation. Related data were represented as *prior* and *posterior* probabilities. Prior probabilities were obtained by calculating the proportions of correct answers falling in each of the four memory awareness categories for each participant. For posterior probabilities, the proportion of correct answers from the total of answers given in each memory awareness category was computed individually for each participant. All participants across the five conditions completed the same spatial memory task a week after the initial experiment reporting on memory recall, confidence and awareness states. Their memory recall scores indicated the amount of objects (out of 21) that they remembered accurately (out of three choices of box, sphere and pyramid) for the specified positions of the diagrams given (Figures 5.5-5.8).

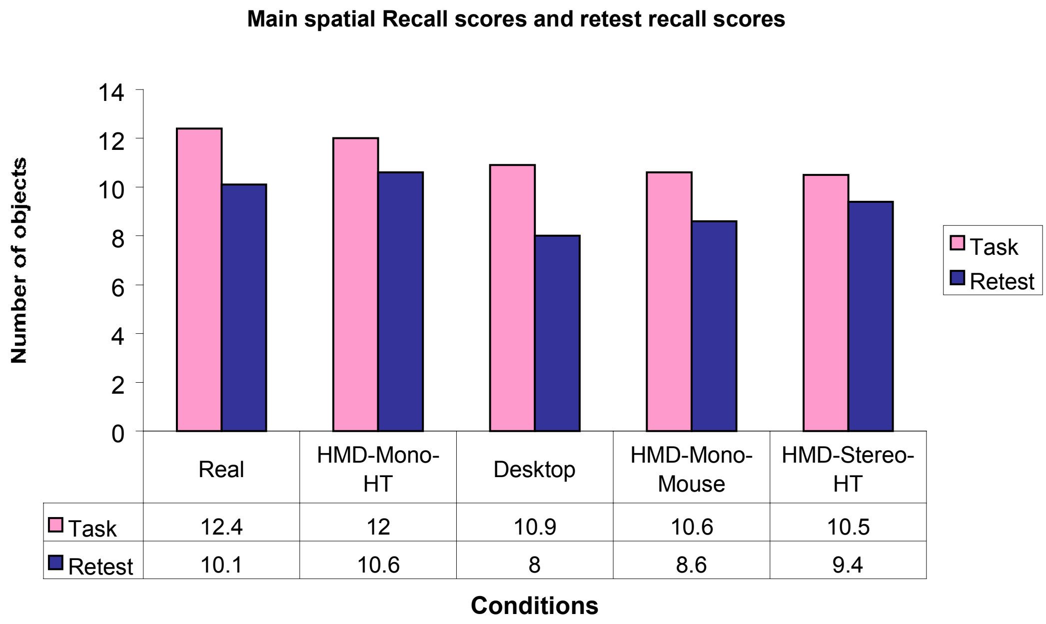


Figure 5.13: Mean correct recall scores (number or correctly recalled objects after exposure to the environment) for the initial spatial task and retest across conditions.

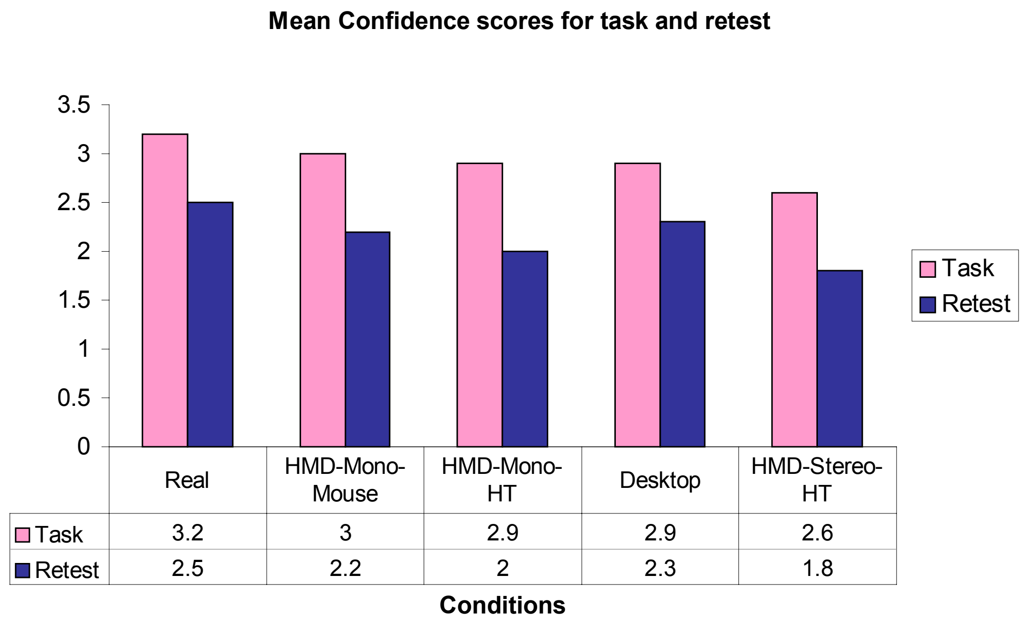


Figure 5.14: Mean Confidence scores for task and retest across conditions (5-point scale).

Figure 5.13 shows the mean accurate recall scores for the spatial memory task as well as for the retest of the same task one week after the initial experiment. For the main task, the mean score of accurate recall was higher for the real-world condition but overall, the ANOVA did not reveal a significant effect across conditions. Similarly, for the retest the ANOVA did not show any significant differences for accurate recall across conditions. However, accurate spatial recall scores were significantly higher for the initial task for certain conditions compared to the equivalent overall recall scores of the retest as analysed for each condition as follows:

- Real-world condition: $F(1,41)=2.765$, $p>0.05$
- HMD mono head tracked condition: $F(1,41)=0.811$, $p>0.05$
- HMD stereo head tracked condition: $F(1,41)=1.197$, $p>0.05$
- HMD mono mouse condition: $F(1,41)=6.308$, $p<0.05$
- Desktop condition: $F(1,41)=4.503$, $p<0.05$

These results show that there was no significant difference in accurate recall between the initial task and the retest for the real-world and the two HMD head tracked conditions. However, the accurate memory recall scores for the conditions utilising a mouse interface (HMD mono mouse and desktop conditions) were significantly lower in the retest.

Figure 5.14 shows the mean confidence scores for participants' responses during the initial spatial memory task and the retest. No significant difference was revealed for confidence scores across conditions globally for the initial task and the retest. There is a tendency for confidence to be significantly higher for the real-world compared to the HMD stereo head tracked condition, but this is a rather weak effect ($p<0.09$). However, confidence scores were significantly higher for the initial task compared to the equivalent confidence scores of the retest for all conditions as follows:

- Real-world condition: $F(1,41)=8.687$, $p<0.01$
- HMD mono head tracked condition: $F(1,41)=10.991$, $p<0.01$
- HMD stereo head tracked condition: $F(1,41)=10.058$, $p<0.01$
- HMD mono mouse condition: $F(1,41)=16.363$, $p<0.01$
- Desktop condition: $F(1,41)=12.653$, $p<0.01$

The experimental design for the main study presented in this chapter was purposefully formal. The FoV was restricted in the real-world condition to match the FoV of the displays. The computer graphics rendering was computed taking into account real world photometric measurements resulting in a photorealistic rendering as described in the previous section. The visual 'similarity' was proven to be the most significant element for the successful completion of the task since no significant difference was revealed in terms of accurate recall, globally for the initial task and the retest across conditions. The same was true for the monocular vs. stereo computer graphics imagery displayed on the HMD conditions. Most importantly, however, the navigational interface influenced memory recall retained in time. It would be of interest to investigate whether a lower quality rendering would have made an impact on spatial memory recall. In the Dihn et al. spatial memory study [DWH99], however, lower resolution of textures did not have an impact on the actual memory recall.

Figure 5.15 shows the mean prior probabilities for the four awareness states across the five conditions. Prior probabilities represent the probability that a correct memory recall response is attached to a specific memory awareness state and indicate the proportion of correct answers for each of the memory awareness states. The memory awareness data were analysed using ANOVA and Tukey's Post-Hoc tests. Generally, there was a significant main effect of condition upon the 'remember' awareness state, $F(4,104)=3.016$, $p<0.05$, and a tendency of significance for the 'know' awareness state, $F(4,104)=1.913$, $p<0.1$. In particular, the probability that correct responses would fall under the 'remember' awareness state was significantly higher for the HMD mono mouse condition compared to the HMD mono head tracked and HMD stereo head tracked conditions ($p<0.05$). This confirms a similar result in the preliminary study presented in Chapter 4 where the proportion of correct responses for the 'remember' awareness state was significantly higher for the HMD mono mouse condition compared with the real-world condition and the desktop condition. Although, this result was not replicated as such in this study, the fact that a significantly higher amount of correct 'remember' responses was revealed for the HMD mono mouse condition in this case as well can be explained as follows:

Generally, the way the spatial memory task was designed allowed for certain mnemonics' strategies to be developed by the participants in each condition. The duration of the exposure to the environment (3 minutes) was long enough for the participants to develop strategies of recollection. For instance, it was observed that a portion of the participants did not recall the positioning of the objects visually, but by memorising patterns of words representing those objects ('box', 'sphere', 'pyramid') according to their positions. The participants that employed a mnemonics' strategy based on words instead of trying to mentally retain elements of the space visually reported the 'know' awareness state when they were making that specific recollection. This resulted in a high proportion of correct responses under the 'know' awareness state. Even if it is assumed that only a proportion of the 'know' responses involved word-based mnemonics, the 'know' awareness state, by definition, is not linked with visual recollections. The participants who mentally visualised the room and the objects in the room during retrieval had a higher proportion of correct responses under the 'remember' awareness state.

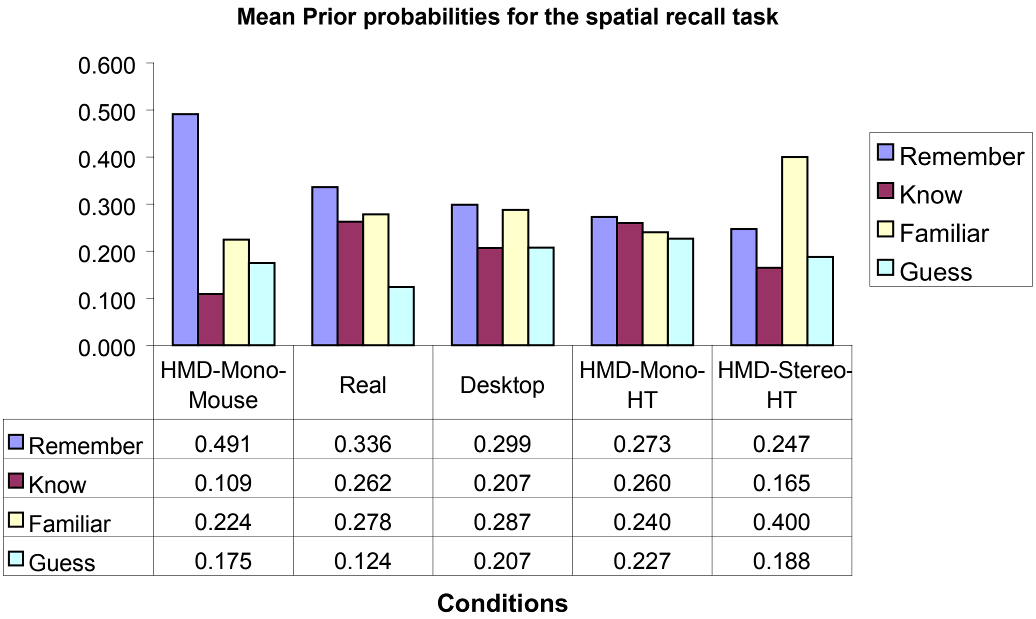
Generally, mnemonics' strategies are common in traditional memory research. There are certain ways of controlling their development, for instance, allowing for shorter exposure times to the stimulus. In this study such strategies of recall were not perceived as a problem since by incorporating the memory semantics methodology, the actual recollections achieved through word mnemonics were assigned to the 'know' awareness state. This methodology offers an insight into 'how' a recollection is achieved during retrieval as well as the relevant distributions between the probabilities associated with the awareness states, even if there are not any significant differences for memory awareness results across conditions. If relevant results show that in certain conditions a weaker trend of non-visually induced recollections is employed by participants towards stronger visually induced recollections expressed by the 'remember' awareness state, it could be assumed that this particular condition affected their mental representation of a space and involved more 'vivid' recollections. Participants were selecting the 'know' awareness state when their recollection was not visual as well as when they employed word-based mnemonics.

In the preliminary study of Chapter 4 it was shown that the correct 'know' responses were minimal. The questions related to that memory task were associated with the space itself and certain meaningful objects in that space and not with artificially placed objects as the primitives in this main study. For the task presented in the main study the responses under the 'remember' and 'know' categories, in particular, reveal the varied mental strategies that participants employed to complete the memory task during retrieval for high-confidence recollections. The HMD mono mouse condition could be perceived as the most ‘unnatural’ condition, since there is no head-tracking interface involved, therefore, no proprioception cues are available. There is a variation of the responses associated with this condition compared to the head-tracked conditions mostly relevant to the amount of correct ‘remember’ responses, signifying a variation in the mental representations of the space and its elements after the exposure to the environment.

Figure 5.16 shows the mean posterior probabilities for the four awareness states across the five conditions. Posterior probabilities represent the probability that a memory recall response assigned to each of the memory awareness states is accurate. An ANOVA applied to those probabilities did not reveal a significant effect of condition. A higher probability that responses assigned to the ‘remember’ and ‘know’ awareness state would be correct, was revealed across conditions.

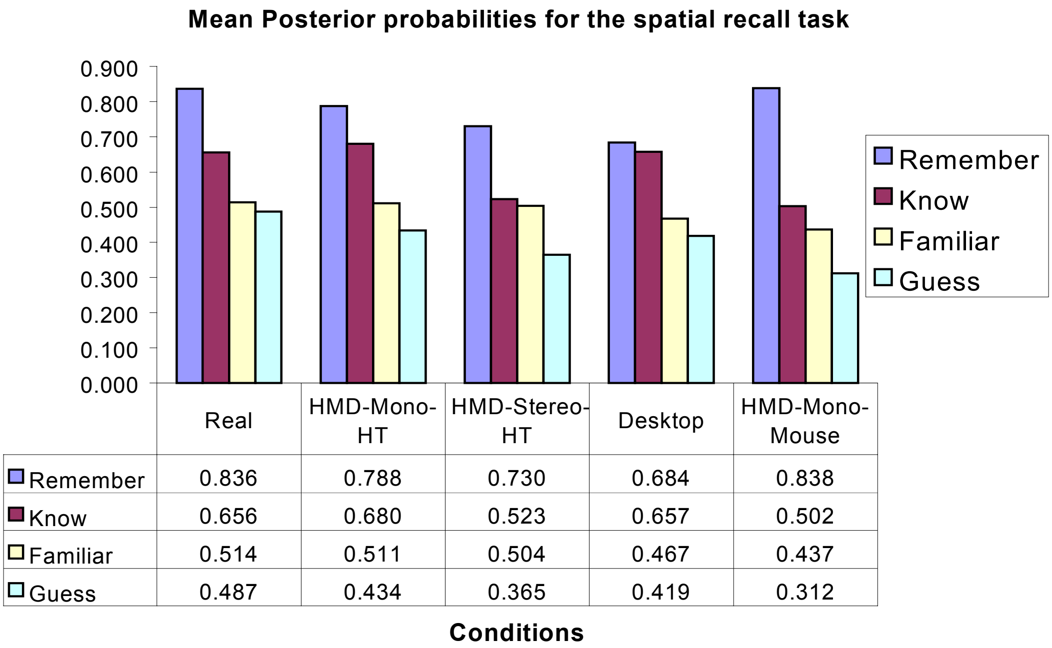
The prior probabilities analysis for the retest reveals no significant effect on any of the four awareness states across conditions (Figure 5.17). It is worth noting, though, that there was a large reduction in the proportion of correct ‘remember’ responses in the initial memory test followed by a significant rise in the proportions of correct ‘guess’ responses for the retest.

Figure 5.18 shows the posterior probabilities across conditions for the retest, only for the ‘familiar’ and ‘guess’ awareness state. A small number of participants assigning their (correct or incorrect) responses to the ‘remember’ and ‘know’ awareness states resulted in posterior probabilities not being calculated reliably. Most of the participants’ responses were assigned to the ‘familiar’ and ‘guess’ awareness states for the retest. A high probability that the responses assigned to the ‘familiar’ and ‘guess’ awareness state would be correct was revealed across all conditions.



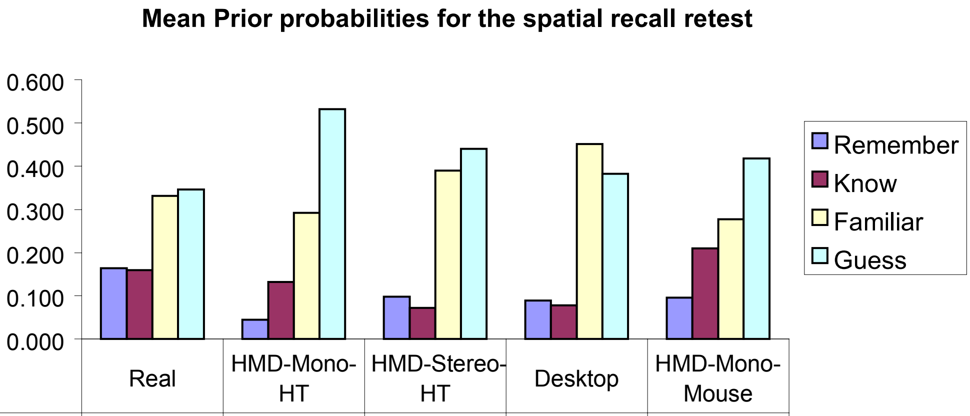
	HMD-Mono-Mouse	Real	Desktop	HMD-Mono-HT	HMD-Stereo-HT
Remember	0.491	0.336	0.299	0.273	0.247
Know	0.109	0.262	0.207	0.260	0.165
Familiar	0.224	0.278	0.287	0.240	0.400
Guess	0.175	0.124	0.207	0.227	0.188

Figure 5.15: Mean Prior probabilities for the initial spatial recall task across conditions.



	Real	HMD-Mono-HT	HMD-Stereo-HT	Desktop	HMD-Mono-Mouse
Remember	0.836	0.788	0.730	0.684	0.838
Know	0.656	0.680	0.523	0.657	0.502
Familiar	0.514	0.511	0.504	0.467	0.437
Guess	0.487	0.434	0.365	0.419	0.312

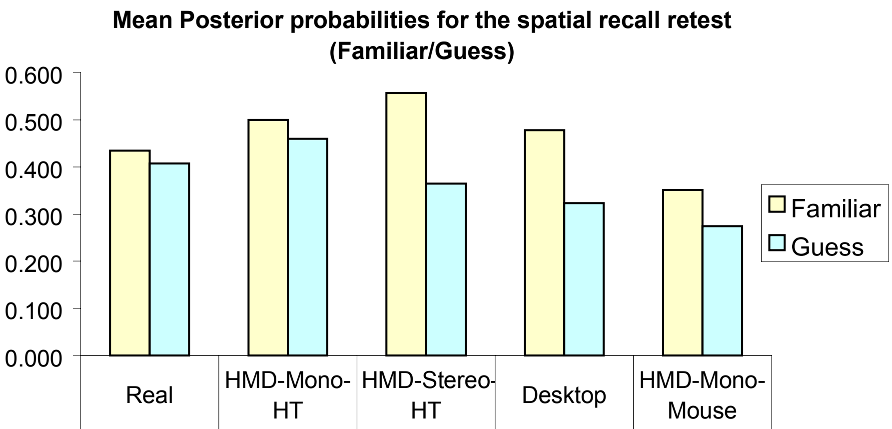
Figure 5.16: Mean Posterior probabilities for the initial spatial recall task.



	Real	HMD-Mono-HT	HMD-Stereo-HT	Desktop	HMD-Mono-Mouse
Remember	0.164	0.044	0.098	0.089	0.096
Know	0.159	0.132	0.072	0.078	0.210
Familiar	0.331	0.292	0.390	0.451	0.277
Guess	0.346	0.532	0.440	0.382	0.417

Conditions

Figure 5.17: Mean Prior probabilities for the spatial recall retest.



	Real	HMD-Mono-HT	HMD-Stereo-HT	Desktop	HMD-Mono-Mouse
Familiar	0.434	0.500	0.556	0.478	0.352
Guess	0.407	0.459	0.365	0.324	0.275

Conditions

Figure 5.18 Mean Posterior probabilities for the spatial recall retest (Familiar/Guess).

	Real	HMD Mono HT	HMD Stereo HT	HMD Mono Mouse	Desktop
Remember	$F(1,41)=4.825$ $p<0.05$	$F(1,41)=10.553$ $p<0.05$	$F(1,41)=11.181$ $p<0.05$	$F(1,41)=40.273$ $p<0.001$	$F(1,41)=4.536$ $p<0.05$
Know	$F(1,41)=1.807$ $p>0.05$	$F(1,41)=2.305$ $p>0.05$	$F(1,41)=2.719$ $p>0.05$	$F(1,41)=2.092$ $p>0.05$	$F(1,41)=5.931$ $p<0.05$
Familiar	$F(1,41)=0.431$ $p>0.05$	$F(1,41)=0.651$ $p>0.05$	$F(1,41)=0.101$ $p>0.05$	$F(1,41)=0.069$ $p>0.05$	$F(1,41)=6.033$ $p<0.05$
Guess	$F(1,41)=9.176$ $p<0.05$	$F(1,41)=15.072$ $p<0.001$	$F(1,41)=11.876$ $p<0.05$	$F(1,41)=9.547$ $p<0.05$	$F(1,41)=5.308$ $p<0.05$

Table 5.2: ANOVA results after comparing test and retest prior probabilities for each of the memory awareness states ($p<0.05$ indicates statistical significance).

The amount of correct 'know' and 'familiar' responses remained almost unchanged across all conditions except the desktop condition where the shift between the 'remember' and 'know' responses towards correct 'familiar' and 'guess' responses was, significantly, the largest. Table 5.2 confirms the above findings by showing the ANOVA results after comparing the prior probabilities acquired from the initial task and retest scores, for each of the memory awareness states across conditions. Although there were no differences of accurate recollections between the initial task and the retest for the real-world, HMD mono head tracked and HMD stereo head tracked conditions, it is noted that the retest recall scores were achieved by following cognitive strategies relevant to the 'familiar' and 'guess' awareness states. The shift between 'remember' responses to 'guess' responses did not signify any difference in the actual amount of correct recollections for those conditions. However, this was not true for the desktop and the HMD mono mouse conditions (both utilising the mouse as the interaction device). The same shift signified a significantly lower amount of recollections for the retest compared to the initial task, for those conditions. Accurate recall scores or even confidence scores could not have revealed such a result in this form which could have implications when certain applications require particular cognitive strategies of spatial perception.

Generally, incorporating awareness states in a memory test connects memory recall with the cognitive process that participants employ and forms a framework towards investigating 'how' humans achieve a task from a cognitive point of view rather than a task performance point of view. Most importantly, fidelity assessments for computer graphics simulations, therefore, could be based on such metrics which could form an integral part of the commonly employed performance efficiency.

5.2.2 Presence Results

The presence questionnaire was administered to the five groups assigned to the real, the HMD mono head tracked, the HMD stereo head tracked, the HMD mono mouse and the desktop conditions. Presence data were analysed using a comparison of means before carrying out an **AN**alysis **O**f **VA**riance (ANOVA) across conditions. In addition to this generic analysis and to avoid the theoretical problem of ordinal data, a binomial regression analysis was employed based on the count of high scores out of six presence questions and following the analysis explained in the Slater et al. study [SSMM98]. The binomial regression was accomplished by assigning 0 if the count of high scores was 0-2 and 1 if the count of high scores was 3-6. Binomial regression, generally, shows the probability of falling under one of the 0 or 1 binomial distributions for each group (condition). This method verified the results related to significant differences acquired from the generic ANOVA statistics.

An overall effect of condition was not revealed for the perceived presence dataset ($F(4,104)=1.685$, $p>0.05$). The overall means for presence are shown in Figure 5.19. There was a significant overall effect of condition for two out of the six presence questions with perceived presence being higher for the real-world condition compared to the technological conditions. The full questionnaire can be found as used in this study in the Appendix A2. To calculate the presence scores, questions 4,5,7,9,11,13 out of this questionnaire were utilised. The overall effect for each of the six presence questions is as follows (Figure 5.20):

Question 4: $F(4,104)=1.647$, $p>0.05$
Question 5: $F(4,104)=5.231$, $p<0.001$
Question 7: $F(4,104)=0.142$, $p>0.05$
Question 9: $F(4,104)=3.603$, $p>0.05$
Question 11: $F(4,104)=2.852$, $p>0.05$
Question 13: $F(4,104)=3.275$, $p<0.05$

For Question 2 (‘To what extent were there times during the experience when the seminar room was the reality for you?’), post-hoc Scheffé tests indicated that the level of presence was significantly higher for the real-world condition compared to each of the technological conditions ($p<0.05$). No significant difference was revealed across the four technological conditions for that particular question. For Questions 6 (‘During the time of the experience, did you often think to yourself that you were actually in the seminar room?’), perceived presence was significantly higher only for the real-world condition compared to the HMD mono head tracked condition ($p<0.05$) with a tendency towards significance compared to the HMD stereo head tracked condition for that question ($p<0.09$). It has to be noted that the actual mean value for presence in the real-world condition, where the FoV was restricted by the appropriate goggles to match the FoV of the HMD and desktop displays, was not as high as the mean value for perceived presence in the preliminary study of Chapter 4 where the FoV for the real-world condition was not restricted. There was no significant effect for presence across the technological conditions; a similar outcome as in the preliminary study. These results also confirm similar effects of condition on presence in studies where the validity of the questionnaire is examined [UCAS2000].

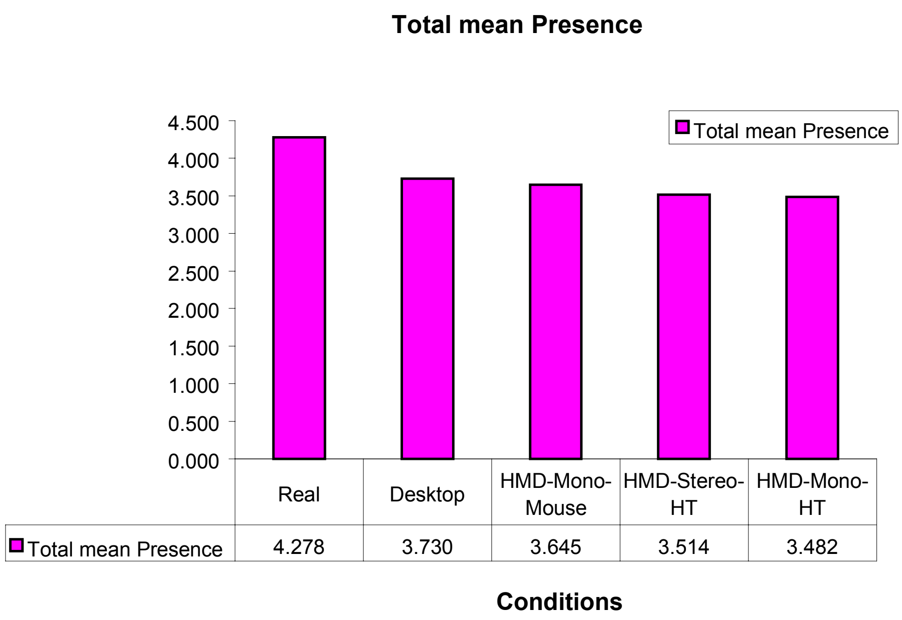
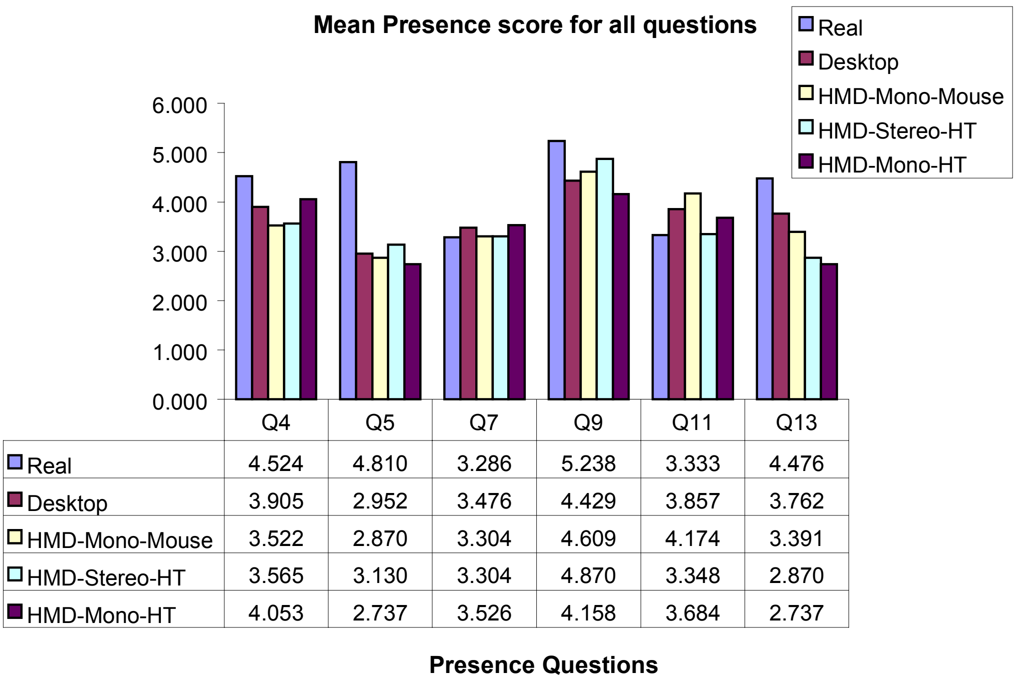


Figure 5.19: Total mean Presence scores across conditions.

Again, in this study one wonders if the questionnaire employed failed to pick up the difference across conditions or if there was not any difference across conditions in the first place, due for instance to the high quality of the rendering or the equivalence of the FoV across conditions. The latter could be the case. Only a concrete understanding of presence, in a way that will allow formal assessments of its perceived level in experimental studies such as this one (if this is ever possible or desirable) will aid towards forming relevant conclusions.



	Q4	Q5	Q7	Q9	Q11	Q13
Real	4.524	4.810	3.286	5.238	3.333	4.476
Desktop	3.905	2.952	3.476	4.429	3.857	3.762
HMD-Mono-Mouse	3.522	2.870	3.304	4.609	4.174	3.391
HMD-Stereo-HT	3.565	3.130	3.304	4.870	3.348	2.870
HMD-Mono-HT	4.053	2.737	3.526	4.158	3.684	2.737

Figure 5.20: Mean levels of reported Presence for all questions.

5.2.3 Subjective Responses to Lighting and Aftereffects Results

The subjective responses to lighting results did not reveal any significant differences across conditions (Figure 5.21). This is not a surprising result since the computer graphics rendering (and the illumination) was similar across the technological conditions [Man2001b]. The photometry measurements acquired from the real world space also ensured that illumination was simulated as accurately as possible between the real world and the computer graphics rendering. A significant negative correlation was revealed, however, between the subjective responses to lighting dataset and the presence dataset for the HMD mono head tracked condition (r=-0.47, Pearson's

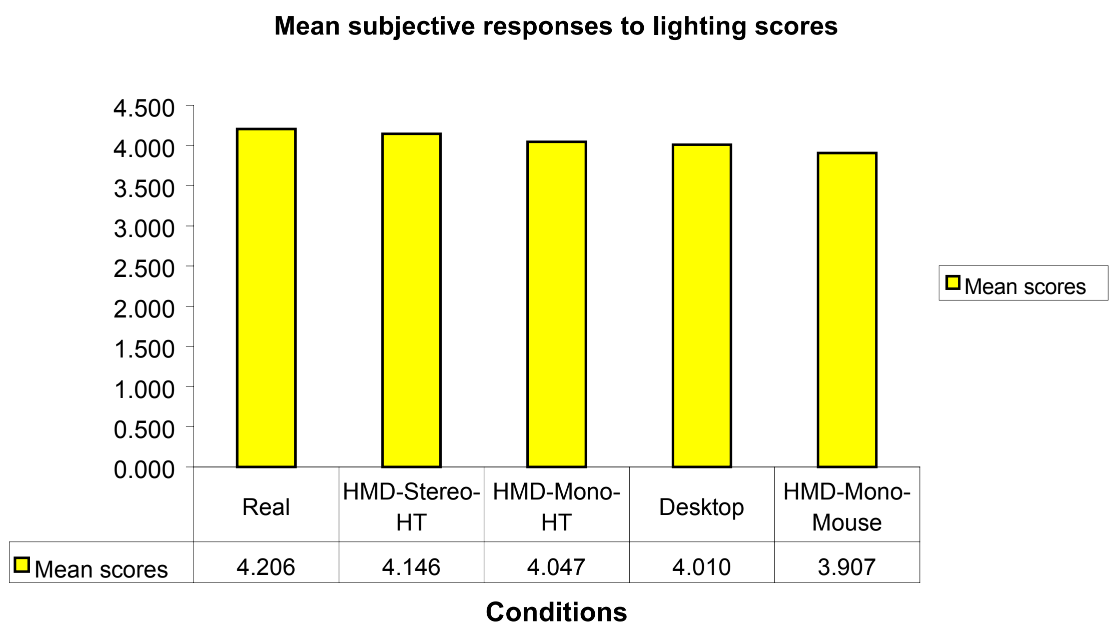


Figure 5.21: Mean subjective responses to lighting scores across conditions.

correlation, $p<0.05$; Spearman's correlation, $p<0.05$) and for the HMD mono mouse condition (r=-0.37, Pearson's correlation, $p<0.1$; Spearman's correlation, $p<0.05$). A negative correlation suggests that if one variable tends to increase, the other decreases (co-varies). Generally, strength (r) is a measure of the correlation but significance assesses how unlikely such a correlation was to occur under the null hypothesis (usually that the population correlation is zero) [Coo99]. Pearson's correlation is based on deviations from the means for each group of data. Spearman's correlation looks at the rank a participant receives for their score on one of the two variables and compares this with the rank they received for their score on the other variable [Coo99]. Interestingly, according to these correlations a high level of perceived presence resulted in a significantly high rating of 'comfort' or 'pleasant' feeling associated with subjective responses to lighting. The subjective responses to lighting questionnaire can be found in Appendix A2. Ratings are structured in a somewhat reverse way to the presence questionnaire since high ratings for presence indicate a high level of perceived presence, however, low assessments for subjective responses to lighting results in a 'radiant', 'warm', or 'pleasant' response. For this phenomenon to be verified and fully explained, a study that would focus on validating these results should be designed.

The Simulator Sickness Questionnaire (SSQ) scores across conditions were quite low and there was not a significant effect of condition for the pre-exposure scores ($F(4,104) = 1.252$, $p>0.05$) and the post-exposure scores ($F(4,104)=0.818$, $p>0.05$). The overall differences between the SSQ scores before and after exposure were also minimal; the three minutes exposure time was too small to provoke severe aftereffects. No significant correlations were revealed by the SSQ scores with perceived presence levels across conditions except for a slight tendency towards a significant negative correlation for the HMD mono head tracked condition. An equivalent negative correlation between perceived presence and SSQ scores was also revealed in the preliminary study of Chapter 4, however, this negative relationship has been supported as well as contradicted in literature as described in detail in Chapter 2.

5.2.4 Digital Monitoring of Navigation Results

In studies where interfaces such as head tracking are incorporated in experimental designs which include more traditional interfaces such as the mouse, unravelling the effect of the actual navigational interface on task performance and awareness states is important. The difference of interaction interface between conditions should be taken into account. Monitoring of participants' navigational patterns provides an aid towards this direction. Participants across conditions could be video–recorded (in the real-world or during their interactions with a synthetic world on the screen) during exposure, allowing in most cases for a qualitative analysis of navigation. A numerical analysis, though, makes monitoring more formal and subject to detailed statistical analysis. This does not mean that qualitative analysis in general, e.g. post–interviews of participants after an experiment should not be carried out. Formal monitoring, however, offers a complete picture of participants' interactions.

In the studies presented in this chapter, participants across conditions were monitored in terms of their navigation patterns while horizontally rotating around their viewpoint, placed in the centre of the experimental room. In particular, participants in the real-world, HMD mono head tracked and HMD stereo head tracked conditions were monitored by means of a digital compass, firmly attached to the back of the swivel chair they were sitting on. This was a wireless device and two angle positions were acquired for each second from 0 to 360. More specifically, direction readings were obtained with the 2-axis electronic compass utilising magneto-inductive technology. This was connected to a wristop PC via a PIC microcontroller interface enabling readings to be recorded at a rate of 1Hz. (Figure 5.10). If the participants were not moving, the same (or largely similar) angle position number was stored indicating idle time. The participants in the desktop and HMD mono mouse condition were monitored by means of software following mouse movements. The participants in these conditions were not rotating the swivel chair they were sitting on but navigated the scene with a common mouse. One angle value from 0 to 360 was acquired per frame. The statistical analysis of this data was based on the amount of idle time. Idle time could provide a means of understanding participants' navigational behaviour while completing the task. Idle time could also offer assessments regarding the level of ease of use of the interface and an indirect measure of overall 'movement' or amount of interactions for each participant. The less idle time participants utilised, the more they navigated around the experimental space (real or computer graphics) and that could have an effect on their overall performance and memory awareness states.

An ANOVA was carried out as well as detailed correlations between the amount of idle time and accurate memory recall, awareness states, presence, subjective responses to lighting and aftereffects for each participant across conditions. A significant overall main effect was revealed for idle time for the real-world and HMD head tracked conditions (mono and stereo), $F(2, 53)= 5.502$, $p<0.01$. Post-hoc Scheffé tests showed that the amount of idle time for the participants in the HMD stereo head tracked condition was significantly higher than those in the real-world ($p<0.01$) condition. No significant effect was revealed for the desktop compared with the HMD mono mouse conditions, $F(1,41)=2.206$, $p>0.05$. Figure 5.22 shows the mean idle time in seconds for each condition. It has to be noted that comparisons of idle time between the real-world condition or HMD head tracked conditions and the desktop and HMD mono mouse condition are presented here with some caution. There was a substantial difference between these two groups of navigation that should be accounted for. Participants in the conditions with proprioception cues available could navigate the scene by movement of the head even without any movement of the chair.

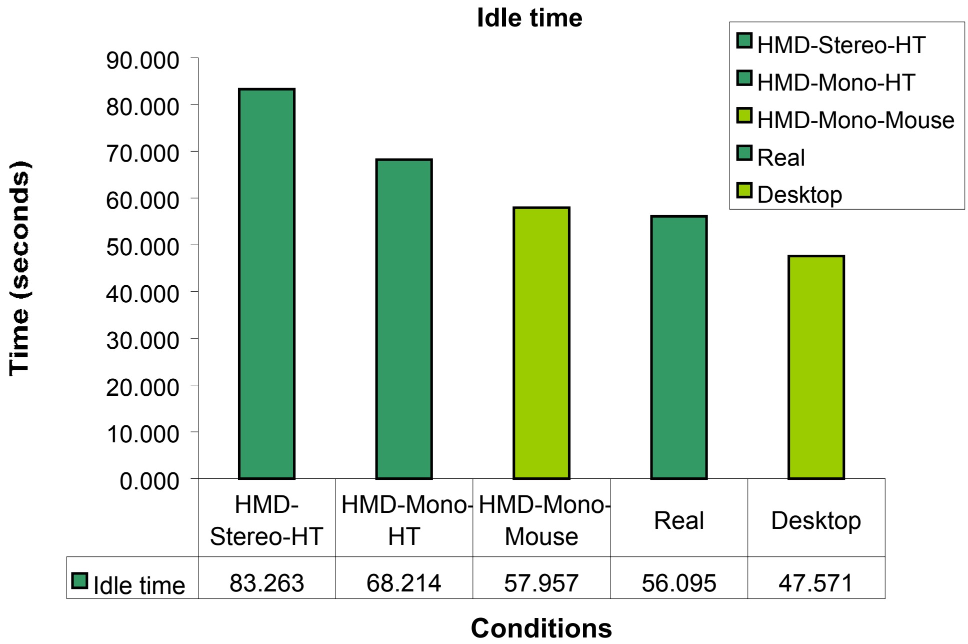
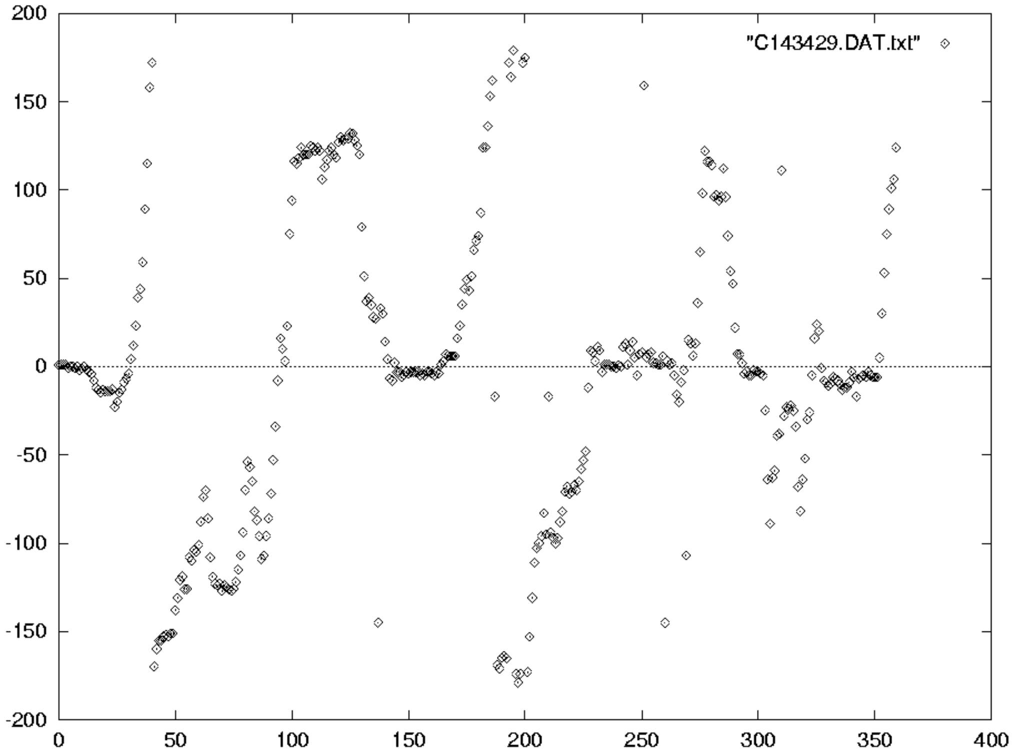


Figure 5.22: Mean idle time across conditions.

This particular action could result in idle time readings. The participants in the conditions without any proprioception cues such as the mouse-related conditions had to change their viewpoint to achieve the same pattern of navigation and this was accounted for as navigation time. It would be, therefore, valid to compare idle time in two separate groups: The real-world and the HMD head tracked conditions (mono and stereo) in one group and the desktop and HMD mono mouse condition in the second group as shown in Figure 5.22. Figures 5.23-5.25 show examples of data stored in the digital compass across the real world and the two head tracked conditions for participants with average idle times. Figures 5.26-5.27 show examples of the navigation data related to tracking mouse movements by means of software for the HMD mono mouse and desktop conditions for participants with an average amount of idle time for each condition.

Correlation analysis between the amount of idle time and task performance, awareness states, confidence level and aftereffects was conducted for each condition. For the real-world, the HMD mono head tracked and the HMD stereo head tracked conditions, no significant correlations were revealed. For the HMD mono mouse condition, a significant positive correlation was revealed between idle time and accurate memory recall of participants ($r=0.42$, Pearson's correlation, $p<0.05$). For the desktop condition, a significant correlation was revealed between idle time and accurate memory recall ($r=0.52$, Pearson's correlation, $p<0.05$), confidence ($r=0.42$, Pearson's correlation, $p<0.05$) and prior probabilities related to the 'know' awareness state ($r=0.45$, Pearson's correlation, $p<0.05$). Interestingly, the above correlations were revealed in the conditions with a mouse interface. A positive correlation indicates that the higher the amount of idle time, the more accurate recollections participants had as well as a higher level of confidence and a higher proportion of correct answers under the 'know' awareness state for the desktop condition. This might mean that in the desktop condition, a higher amount of idle time indicated a higher amount of non-visually induced recollections as opposed to visual mental imagery. This result appears only in that condition and therefore can not be generalised.

Generally, participants in the head tracked conditions utilised a higher amount of idle time viewing the scene during exposure than participants in conditions where the mouse was used for navigation but that could be an effect of the actual monitoring procedure. Idle time indicates not only the amount of time that participants spent relatively still, but also, indirectly, it shows the amount of interaction that participants employed to complete the task. Obviously, a high amount of idle time indicates a low amount of interaction or navigation around the scene. The pattern of navigation, therefore, for the real-world condition is not similar to the HMD stereo head tracked condition. The stereo effect might be the reason why participants spent a significantly higher amount of time being idle in comparison to the real-world condition. It could be argued that the higher amount of idle time for the HMD head tracked condition was a result of participants increased focus due to the stereo imagery. However, there is no correlation between idle time and the amount of perceived presence for that condition. Presence was also not significantly different for that condition compared to the four remaining conditions. Similar significant correlations were not revealed for the real-world, HMD mono head tracked and HMD stereo head tracked conditions.



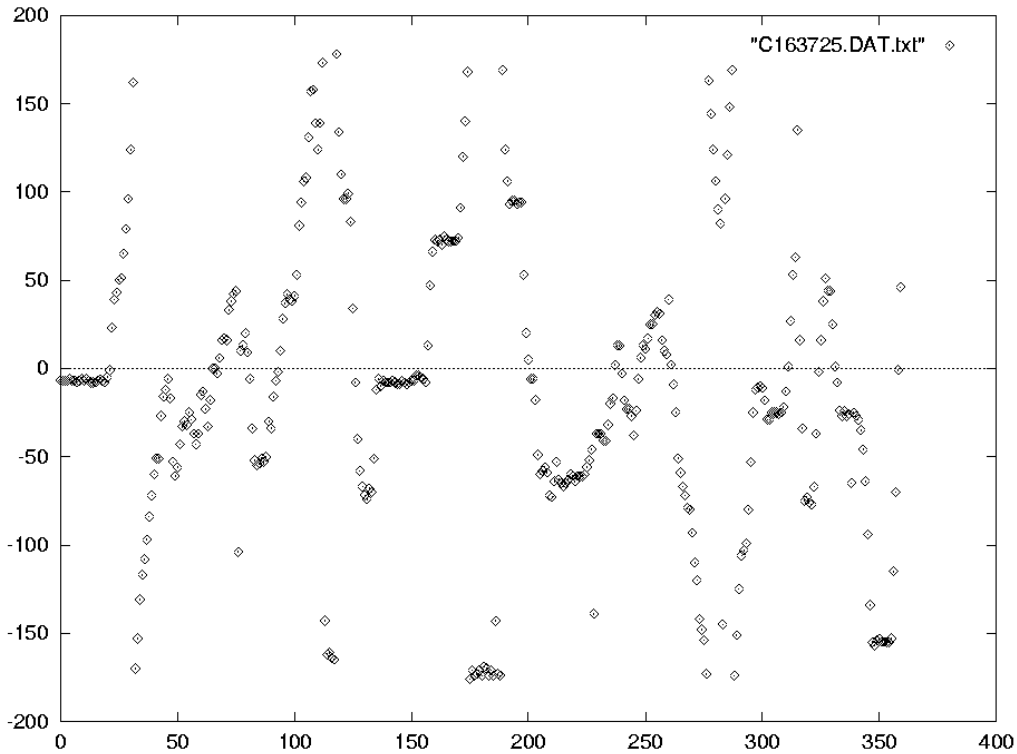
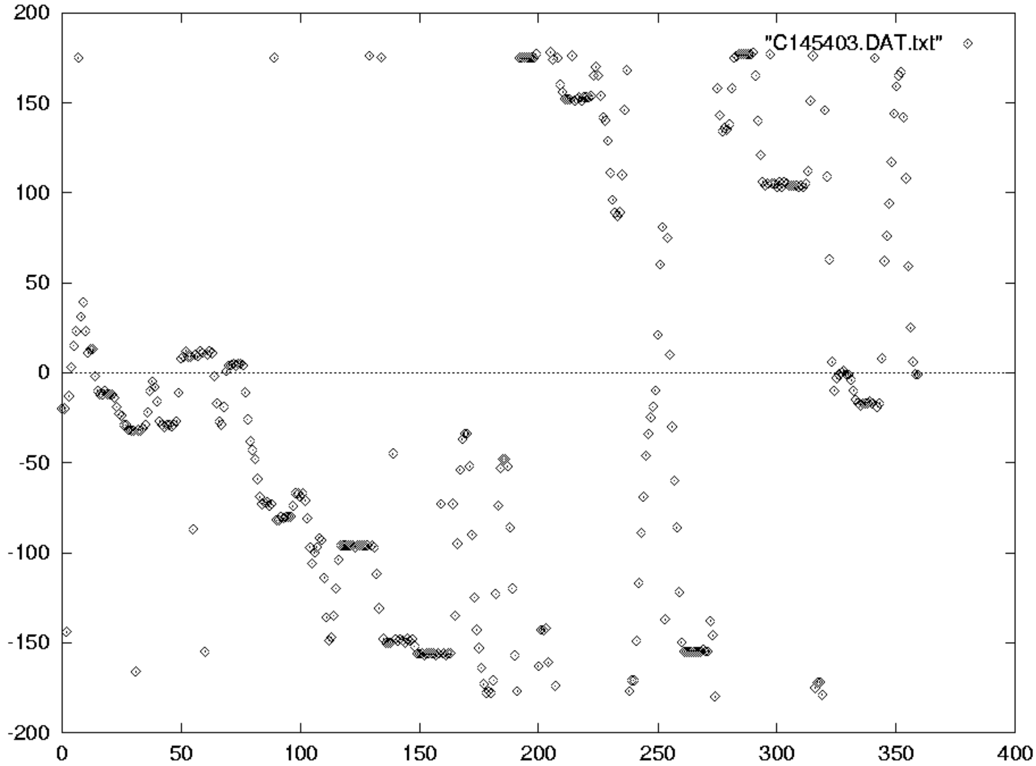


Figure 5.23: Sample graphs for navigation data for the real-world condition with average idle times (x axis is time in half seconds, y axis is angle value in degrees).



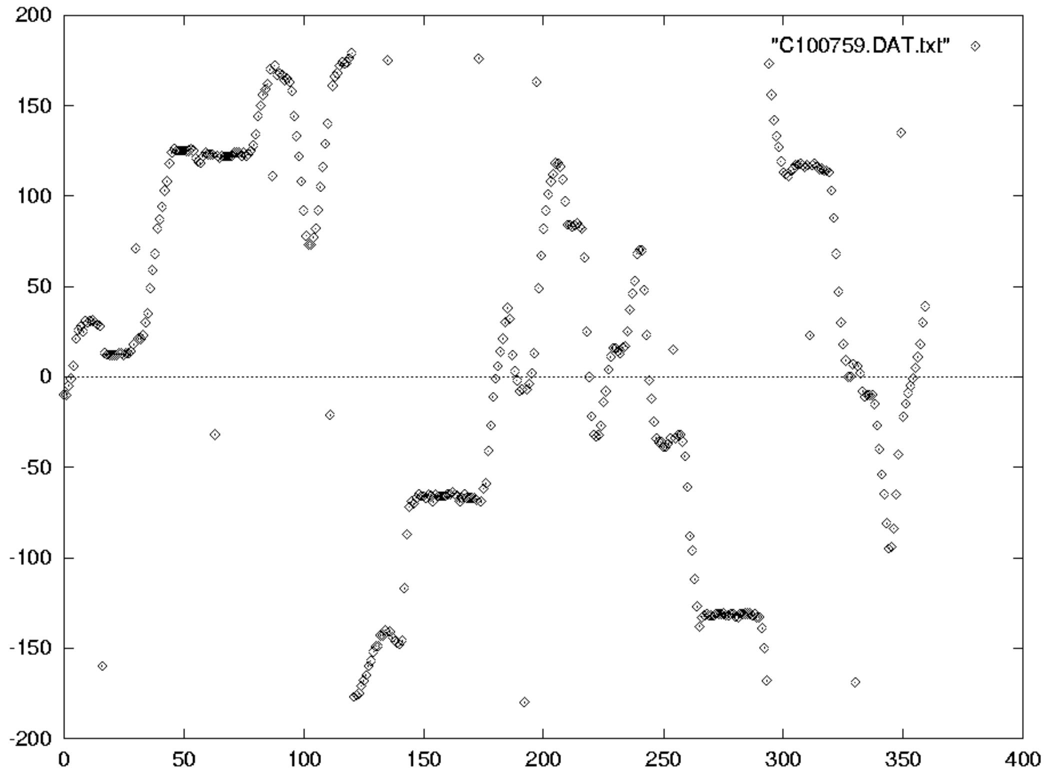


Figure 5.24: Sample graphs for navigation data for the HMD mono head tracked condition with average idle times (x axis is time in half seconds, y axis is angle value in degrees).

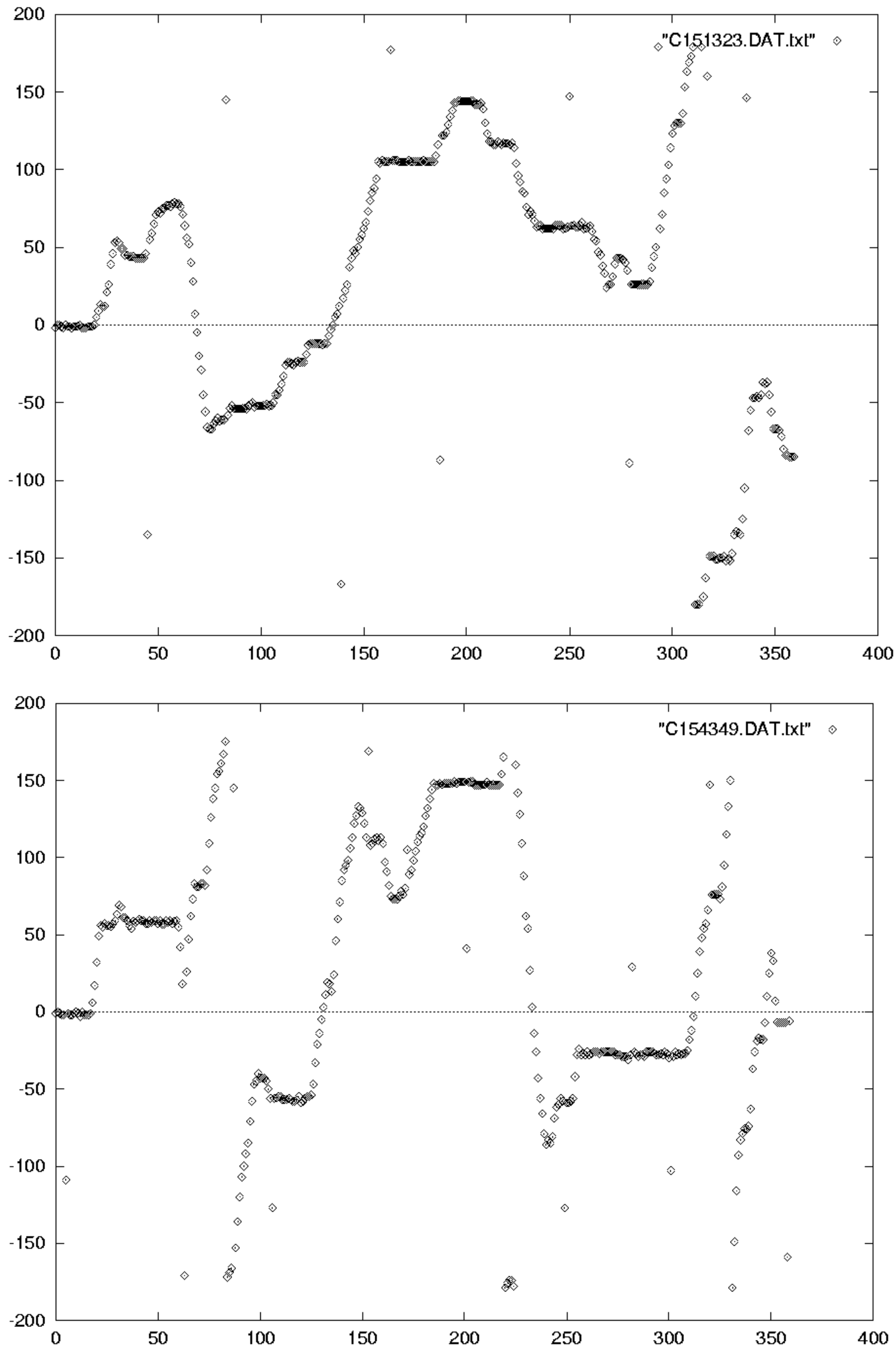


Figure 5.25: Sample graphs for navigation data for the HMD stereo head tracked condition with average idle times (x axis is time in half seconds, y axis is angle value in degrees).

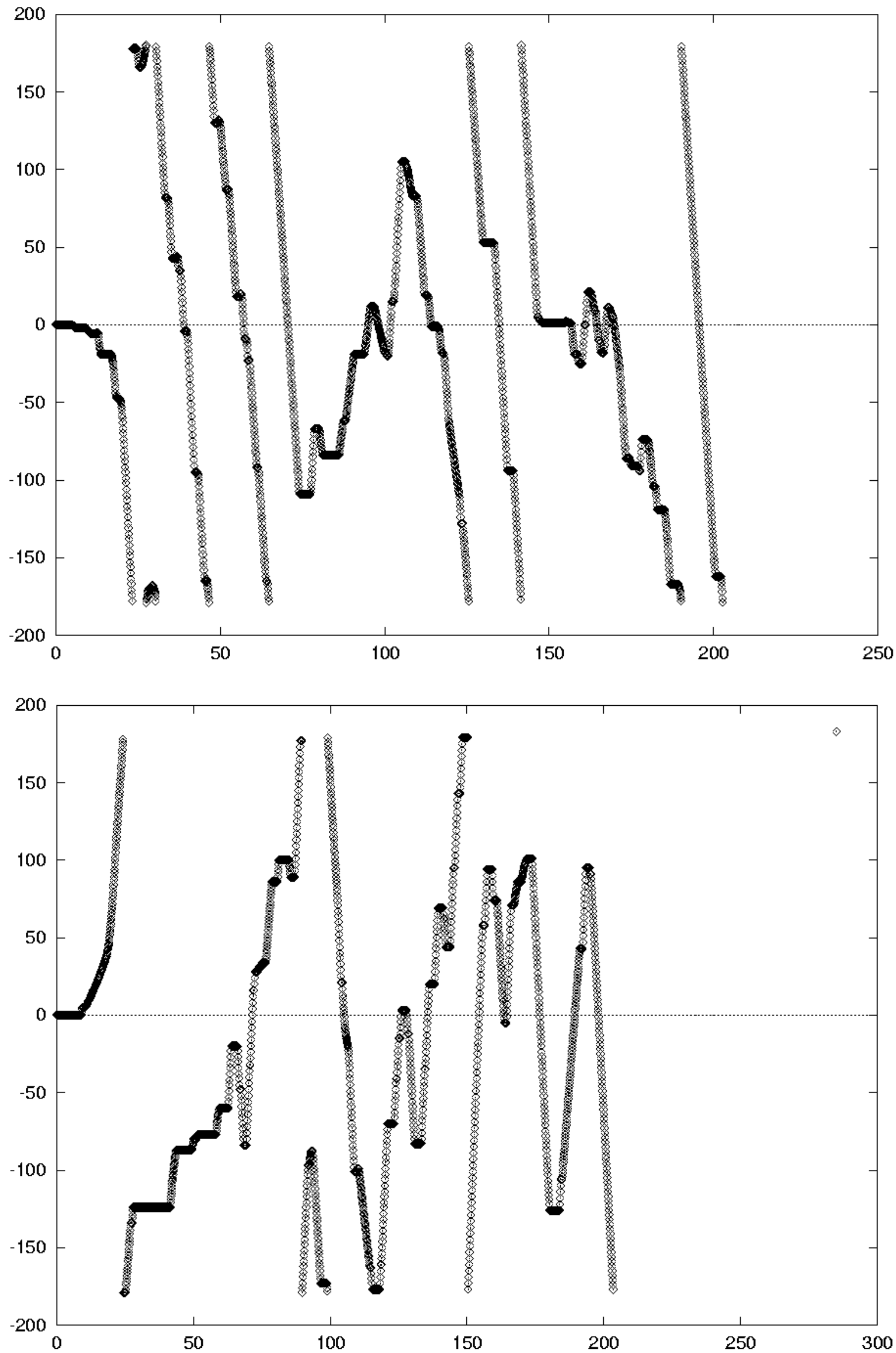


Figure 5.26: Sample graphs for navigation data for the HMD mono mouse condition with average idle times (x axis is time in seconds, y axis is angle value in degrees).

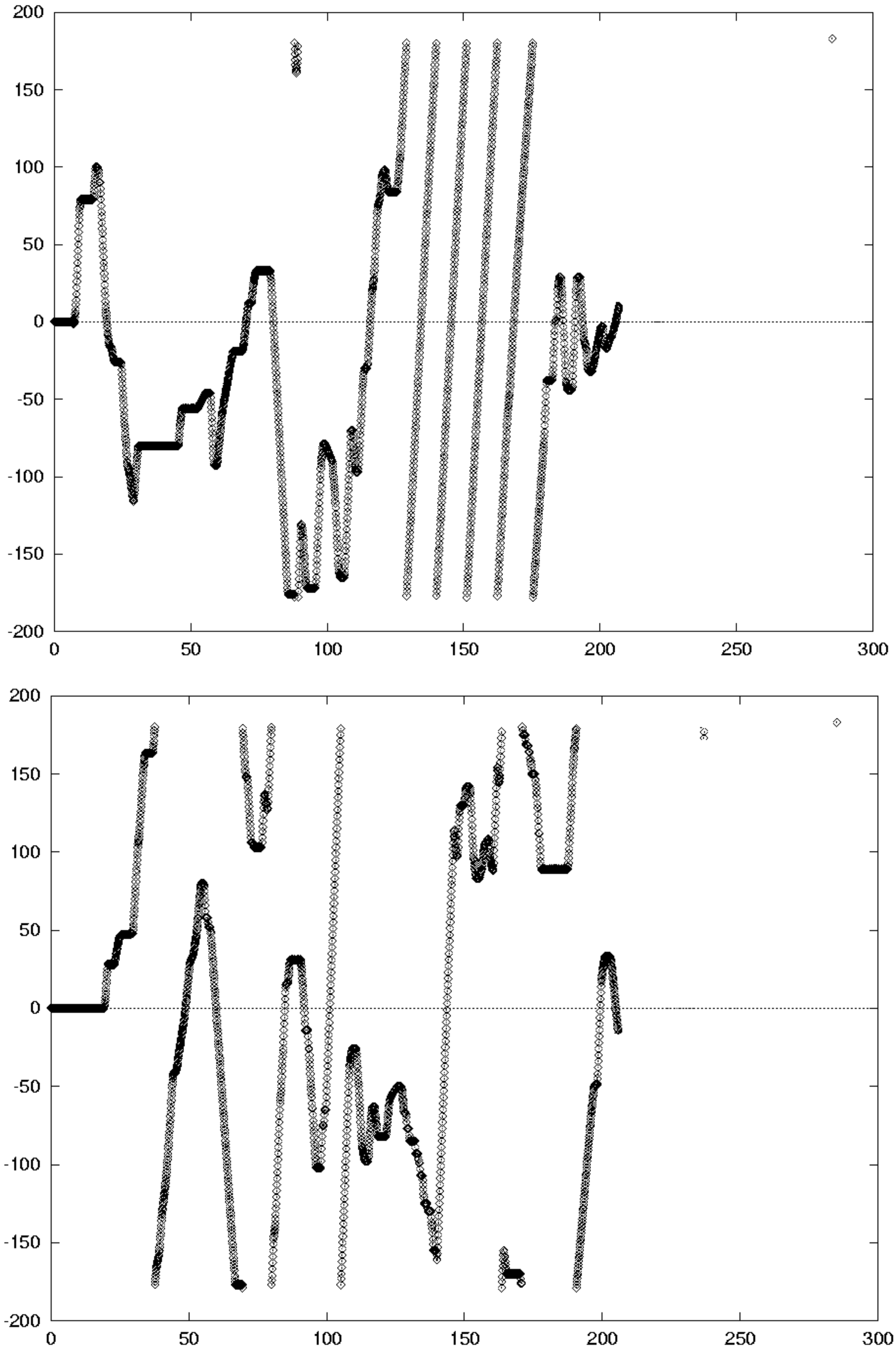


Figure 5.27: Sample graphs for navigation data for the desktop condition with average idle times (x axis is time in seconds, y axis is angle value in degrees).

Generally, the incorporation of such results offers additional information related to participants' behaviour during exposure. To validate the results mentioned here, a more focused study needs to be conducted including more controls such as, for instance, eye tracking for the real-world condition to account for the movements of the head while idle or head tracking monitoring data.

5.2.5 Qualitative Analysis of Participants Comments

The last part of the presence questionnaire included a blank sheet with the following question: 'Please write down any further comments that you wish to make about your experience. What things helped to give you a sense of 'really being' in the space, and what things acted to 'pull you out' of this?' Participants replied to this question directly, chose to report personal feelings and comments about their experience or preferred not to comment at all.

Participants reported the following aspects of the experiment that *enhanced* their feeling of 'being' in the experimental room:

- Realistic rendering and lighting: 10% of participants in the HMD mono head tracked condition, 20% of the HMD stereo head tracked condition, 25% of the desktop condition and 30% of participants in the HMD mono mouse condition.
- Proprioception cues: 40% of participants in the HMD mono head tracked condition and 5% of the HMD stereo head tracked condition.
- The HMD technology: 10% of the HMD mono mouse condition.
- Restriction of FoV (thus blocking the real world): 10% of participants in the desktop condition.

Other issues that, as mentioned, promoted the 'realism' of the experience were the sense of depth for the stereo imagery, the physical room being dark during exposure to the computer graphics environment, the mouse movement emulating head movements, the fact that all the objects in the room could exist in real life and the actual engagement with task completion.

Participants also reported the following aspects of the experiment that *pulled them out* of a feeling of 'being' in the experimental room:

- The HMD allowing for periphery vision of the real space (although the room was darkened, some illumination was shed by the displays): 30% of the participants in the HMD mono conditions, 5% of participants in the HMD stereo head tracked condition.
- General HMD discomfort: 20% of the participants in the HMD mono head tracked condition, 5% of the HMD stereo head tracked condition and 10% of the participants in the HMD mono mouse condition.
- Primitive blue objects not being 'every-day' objects: 20% of the participants in the HMD mono mouse condition, 15% of the desktop and real-world condition, 10% of the HMD stereo head tracked condition and 5% of the participants in the HMD mono head tracked condition.
- 14 frames per second frame rate: 20% of the HMD stereo head tracked condition, 15% of participants in the HMD mono head tracked condition, 15% of the desktop condition and 10% of participants in the HMD mono mouse condition.
- Rendering issues such as textures and resolution: 15% of the participants in the HMD mono head tracked condition and 15% of the desktop condition.
- The mouse (absence of head tracking): 20% of the participants in the HMD mono mouse condition and 10% of the participants in the desktop condition.
- Lag between head movements and update of computer graphics imagery: 15% of participants in the HMD stereo head tracked condition and 10% of the participants in the HMD mono head tracked condition.
- Absence of sound cues: 15% of the participants in the HMD mono mouse condition.
- Monocular graphics imagery: 10% of the participants in the desktop condition and 5% of the participants in the HMD mono mouse and HMD mono head tracked conditions.
- Small FoV of the display compared to natural human vision: 10% of the participants in the HMD stereo head tracked condition and 5% of the participants in the HMD mono mouse and HMD mono head tracked conditions.
- Restricted navigation: 10% of the participants in the HMD mono mouse and real-world conditions and 5% of the participants in the HMD mono and stereo head tracked conditions.

Other relevant issues participants mentioned that pulled them out of the feeling of 'being there' could be attributed to the fact that they were not familiar with HMD technologies, the absence of visible windows in the room and the lack of dirt in the computer graphics room. A participant in the HMD mono mouse condition suggested that computer science students are often critical with computer graphics renderings and this fact could have affected their involvement in the task. This comment is always an issue in experimental studies such as the ones presented in this thesis since, it is common, that participants are usually a subset of the student population of the university where a particular research study is taking place.

Interestingly, certain participants in the HMD conditions reported that after exposure to the computer graphics environment they became disorientated in the real room where the experiment took place in relation to the computer graphics room. This resulted in participants looking for the door, for instance, in the real room at the orientation of the doors in the VE after exposure. A participant also reported that for the first 30 seconds of VE exposure he was aware that he was looking at a computer graphics simulation. After that period of time he became adjusted and perceived the computer graphics space as more or less 'real'.

5.2.6 Observations and Discussion

The set of studies presented in this chapter attempt to assess the simulation fidelity of a VE displayed on a desktop and HMD display, comparing accurate memory recall, awareness' cognitive states and sense of presence of participants against the real world space the VE represents. This is a significant process for the development of such technological interfaces since training in a VE system that would be capable of perfectly simulating the real world should result in the same training effect as the real world.

Spatial perception is essential in the real world. A photorealistic, computer-generated interactive environment for simulation and training strives to achieve the same sense of *space* as in the real world. Subjective measures based on human spatial perception supplementary to accurate geometry, illumination and task performance reveal the actual cognitive mechanisms in the perception of a VE that are not otherwise apparent. Subjective constructs such as the presence construct appeared to address the visual fidelity of the computer graphics application across conditions; an overall effect of condition was not revealed. This could mean that there was no actual difference in presence across conditions, at least, ‘presence’ as communicated by the presence evaluation method utilised. However, the memory awareness states results indicated certain variations relevant to the mental processes participants followed during retrieval in order to complete the task, related to the motor interface. Results also revealed the specific distribution of these mental processes reflected on the memory awareness states for the initial task and the retest. Of greatest importance for the set of two studies presented in this chapter is that the navigation method of a computer graphics world (head movements vs. mouse) has an effect on the cognitive strategy adopted and therefore on the type of mental representation of the scene. In addition, the navigational interface seemed to affect the amount of recollections retained in time since these were significantly lower in the retest for the conditions utilising a mouse interface compared to head tracking or the real world. These results could have important implications for the design of VE systems and applications focusing on the interface for interaction.

Strategies of recall that are based on word mnemonics, generally, are constructed when participants perceive that they will need to remember a sequence of elements of a space to complete a relevant memory task. They are a route straight to a semantic representation bypassing the visual mental representations. So, mnemonics, even though they are the ‘natural’ solution to some recall situations, are only a solution to a small and rather unrepresentative set since there is a low probability that they would occur in normal circumstances. The results in this study show a higher proportion of correct ‘remember’ responses for the HMD mono mouse condition. One wonders if this fact makes this condition ‘realistic’ or of higher simulation fidelity, even if it reveals a larger amount of correct ‘remember’ responses than conditions that incorporate more ‘natural’ motor interfaces such as head tracking. If ‘reality’ is defined by the degree of similarity to the real room, in this case the HMD mono mouse condition is not very ‘real’. However, the cognitive strategies employed for completion of the spatial memory task is affected by the degree of ‘realism’ of the motor response.

Word based mnemonics and generally recollections that were not linked to a specific mental image were identifiable by the high proportion of the correct ‘know’ responses. The utilisation of a viewing method such as the HMD plus an ‘unreal’ motor response such as the mouse, stopped participants employing non-visually induced recollections and resulted in a larger distribution of correct responses assigned to the ‘remember’ awareness state. This distribution observed in the initial task did not affect, though, the amount of accurate recollections in the retest, which was significantly lower for this condition compared to the initial study. By decreasing the degree of ‘reality’ of the motor response, participants -paradoxically- adopt visually induced recollections.

Studies such as the one presented here, therefore, focus on the cognitive strategies of the participants, towards fidelity metrics based on characteristics of displays and input interfaces. Generally, there's little point achieving photorealism for computer graphics rendering if participants employ a strategy to achieve a task that is not similar to its real world counterpart due to the interaction interface employed, when a high level of simulation fidelity is crucial. This is not the case here, however, something less ‘real’, therefore, less computationally expensive but more demanding because of its novelty may restore a more ‘naturalistic’ or desirable cognitive strategy. Research could identify issues as such by using methodologies that allow investigations based on the perceptual processes participants employ in order to achieve a task.

The task employed in this study did not allow free navigation around the experimental space. In that sense, one could argue that the ‘spatial’ element of the task was limiting. The task was chosen, though, with a focus on controlling all variables of the experimental design as well as on applying the memory awareness states methodology in that context. It was crucial to construct a formal experimental design between a real-world situation and its 3D simulation counterpart. Future work could include a task which will provide more opportunities for navigation. By employing methodologies that have been examined and validated through decades of experimentation such as the memory awareness states methodology, computer graphics research and VE technologies can get closer to actually exploiting the human perceptual mechanisms towards successful applications.

5.3 Summary

In this chapter, the main experimental studies (initial task and retest) were presented in detail. The experimental design employed remedied certain shortcomings of the preliminary study analysed in Chapter 4. A visual spatial memory recall task and the memory semantics framework were adopted so as to propose a methodology for assessing the simulation fidelity of a computer graphics scene in relation to its real-world counterpart. The computer graphics rendering took advantage of photometry measurements of the real world room in order to ensure photorealistic quality of the radiosity rendering and simulated illumination that would resemble the real-world illumination. Experimental conditions were expanded compared to the preliminary study to include HMD monocular and stereo graphics imagery as well as head tracking for navigation.

Results for accurate memory recall did not reveal, as predicted, a significant difference across conditions; the same was true for confidence levels of participants across conditions. Since the computer graphics scene was of a high rendering quality, the difference of interface or display device did not affect participants' accurate recall. The amount of accurate recollections between the initial task and the retest significantly differed for certain conditions. More specifically, there was a significantly higher amount of accurate recollections for the real-world condition and the HMD head tracked conditions for the initial task compared to the retest. This was not true for the desktop and HMD mono mouse conditions. Results for memory awareness states revealed a significant overall effect for prior probabilities connected with the 'remember' awareness state as well as a tendency towards significance connected with the 'know' awareness states. A higher proportion of correct 'remember' responses was revealed for the HMD mono mouse condition compared to the HMD head tracked conditions, verifying equivalent results in the preliminary study. The cognitive strategies linked with the recollections in the initial task compared with the retest were significantly different mainly indicating a shift from a 'remember' awareness state in the initial task to a 'guess' awareness state in the retest. It is really important to note that although this shift was observed across all conditions, it did not signify a lower amount of accurate recollections between the initial task and the retest, only for the real and HMD head tracked conditions (mono and stereo). Generally, memory awareness states offered an invaluable contribution towards understanding the strategies that participants followed in order to complete the task and these strategies varied. Task performance by itself, as commonly employed when displays or computer graphics renderings and algorithms are evaluated could not have offered such an insight. This constitutes one of the major contributions of this thesis.

The amount of perceived presence did not prove to be significantly different across conditions. This was true for the level of subjective responses to lighting as well. Certain interesting correlations, however, were revealed between the two datasets for the HMD mono head tracked and HMD mono mouse conditions. According to these correlations, a high level of perceived presence resulted in a high rating of 'comfort' and 'pleasant' feeling associated with subjective responses to lighting The amount of aftereffects was not high since participants' exposure to the computer graphics environment displayed on the HMD was very short (3 minutes).

Monitoring participants' navigation patterns indicated a significant difference between a higher amount of idle time for the HMD stereo head tracked condition compared to the real-world condition. For the HMD mono mouse condition a significant positive correlation was revealed between idle time and accurate memory recall of participants. For the desktop condition, a significant correlation was revealed between idle time and accurate memory recall, confidence and prior probabilities related to the 'know' awareness state. Interestingly, the above correlations were revealed in the conditions including a mouse interface which revealed a significantly lower amount of accurate recollections in the retest, compared to the initial task.

Chapter 6

Conclusions and Future Work

The main scope of this thesis was the development of a methodology or metric that leads to the evaluation of Virtual Environment (VE) implementations with respect to the real-world scenes and task situations they represent. These implementations comprise computer graphics imagery, display technologies such as Head Mounted Displays (HMDs) and interaction interfaces ranging from a mouse to head tracking. This investigation has resulted in an experimental framework that facilitates the comparison between real and computer graphics interactive scenes by using human judgements of spatial memory awareness states, for assessing the fidelity of a computer graphics simulation. Novel experimental results are presented from a ‘cognitive’ rather than ‘task’ point of view.

The ultimate goal of VE technologies for simulation and training is to induce a spatial sense similar to the sense of three dimensional space humans get in the real world. This is essential for applications that often target successful transfer of training in the real world and consequently require a high level of simulation fidelity. Thus, spatial perception issues are clearly involved in interactive computer graphics rendering and should be considered at various levels while assembling a VE implementation. This thesis has showed that experimental studies incorporating human responses could be used in order to evaluate and validate VE technologies with respect to real scenes, thereby, assessing the fidelity of a particular design. The resulting methodology is focused on the perceptual and subsequent cognitive processes involved as opposed to only task performance commonly employed, across experimental conditions.

6.1 Main Contributions

This thesis presents a metric, based on human judgements of spatial memory awareness states for assessing the simulation fidelity of a VE implementation in relation to its real scene counterpart. This framework is based on the cognitive processes or awareness states participants employ in order to retrieve the memory of a space after exposure to a computer graphics simulation representing that space as well as the real world space being simulated. The theory behind the framework is differentiating between retrieval processes and it is drawn from traditional memory research. It offers participants the capability to describe how they make their spatial recollections by selecting between four choices of possible awareness states ('remember', 'know', 'familiar', 'guess'). These depend on the visual mental imagery involved during retrieval or the lack of it, the familiarity of the recollection and also include guesses, even if they were informed. Certain associations to semantic and episodic memory are expressed in Chapter 3. However, since this is quite a controversial issue for traditional memory research, the methodology focuses more on the variations of the mental processes involved in recollection rather than to any associations of this sort. The method also incorporates subjective judgements of the participants' level of presence, responses to lighting and simulator sickness together with monitoring through digital means participants' navigational patterns. This method is criticised against commonly employed frameworks of task performance assessments for computer graphics simulations in Chapter 2 and 3. The main findings of the experimental studies presented in Chapter 4 and 5 are summarised here.

The preliminary study proved the feasibility and applicability of the method revealing a variation of the distribution of participants' awareness states across conditions, especially when task performance failed to reveal any. This preliminary general outcome was later validated in the main studies. It was proved that task performance comparisons, commonly employed and also commonly inadequately designed, are not offering a complete assessment of fidelity if a formal framework is not adopted. The general premise of this thesis is focusing on *'how'* spatial tasks are achieved, rather than only, as in earlier VE research, on *'what'* is achieved.

In particular, for the preliminary study, statistically significant differences were revealed for the 'remember' awareness state linked with mental imagery and event-based recollections and for the 'familiar' awareness state linked with awareness of an event or an image that cannot be accurately placed in time. One of the most interesting results revealed a significantly higher proportion of correct 'remember' responses for the HMD condition compared with the real-world condition. Accurate spatial memory recall, however, was not proved to be significantly different across conditions. The same was true for confidence levels of participants across conditions. As suggested, therefore, memory awareness states' analysis gave an invaluable insight into participants' strategies of retrieval, most interestingly across specific conditions where results for presence and accurate memory recall were not proven to be significantly different.

The main studies verified this premise. In these studies, results for accurate memory recall did not reveal a significant difference across conditions; the same was true for confidence levels of participants across conditions. Since the computer graphics scene for the main studies was of relatively high rendering quality, the difference of the interface (mouse vs. head tracking) or display device (desktop monitor vs. HMD) did not affect participants' accurate recall across conditions, separately analysed for the initial task and the retest. A significant higher amount of accurate recollections was revealed, however, in the initial task for the desktop and HMD mono mouse condition compared with recollections made in the retest for those conditions. Spatial recollections were the same for the initial task and the retest for the real-world and HMD head tracked conditions (mono and stereo). These are the conditions that incorporated proprioception cues.

Results for memory awareness states revealed a significant overall effect for prior probabilities connected with the 'remember' awareness state and a tendency towards significance when connected with the ‘know’ awareness state across conditions for the initial task. This outcome was also revealed in the preliminary study for the same displays and interface condition, in this case, the HMD monocular mouse condition. ‘Remember’ responses indicate a strategy for recollection based on visual mental imagery and this awareness state was generally expected to be expressed in conditions incorporating more ‘naturalistic’ interfaces such as head tracking. This variation did not prove to positively affect the amount of participants’ recollections in the retest of the initial task after one week. It could mean, however, that a virtual interface with the highest level of simulation fidelity according to the interface does not always correspond to the more ‘visual’ spatial perception processes. For applications that do require a high amount of strong recollections based on visual mental imagery, the more natural of interfaces such as head tracking may not be appropriate for this application’s goals. Desirable variations of cognitive strategies for specific application purposes, therefore, could be ultimately identified. It could be true, for instance, that for flight simulation applications it is crucial for trainees to achieve a high level of visually induced recollections related to instruments as opposed to feelings of familiarity of even recollections that are confident but not accompanied by visual imagery.

There was a significant shift from reported accurate responses linked with the ‘remember’ and ‘know’ awareness states in the initial task to accurate responses of ‘familiar’ and ‘guess’ awareness states in the retest across all conditions for the studies presented in Chapter 5. It is really important to note that although this shift was observed across all conditions, it did not signify a difference in the amount of accurate recollections between the initial task and the retest for the real and HMD head tracked conditions (mono and stereo). This does not mean that the amount of accurate recollections was higher for these conditions (it was not). It means that although there was a shift from high-confidence awareness states (‘remember’ and ‘know’) to lower confidence awareness states (‘familiar’ and ‘guess’), this fact did not make a difference in spatial elements of the space retained in time for the conditions that included proprioceptive cues and more ‘natural’ interfaces. Accurate recall was the same between the initial task and retest for those conditions, however, it was lower in the retest for the conditions utilising a mouse even if for all conditions, the shift of cognitive awareness states mentioned above was the same. This is one of the most important results in this thesis.

Generally, matching participants' performance in a simulation of any sort to performance in the real-world situation does not guarantee that the actual cognitive processes of possible simulations that participants employ in order to complete the task would be the same. This fact should be accounted for in the design or the assessment of fidelity since certain of these processes as suggested above, could be desirable for application goals. Task performance scores, therefore, could be taken into account according to the desirable awareness states. For any task other than spatial memory or spatial perception employed in benchmarking processes for VE evaluation, this could be a significant premise.

For the main studies, the amount of perceived presence did not prove to be significantly different across conditions; this was true for the level of subjective responses to lighting as well. A uniform outcome across all the studies presented in this thesis did not reveal any variations in perceived presence for all participants among the technological conditions including traditional displays or any HMD technology. Certain interesting correlations, however, between the two were revealed for the HMD mono head tracked and HMD mono mouse conditions in the main studies of Chapter 5. According to these correlations, a high level of perceived presence resulted in a high rating of 'comfort' or 'pleasant feeling' being associated with subjective responses to lighting. Generally, there is some ambiguity related to the notion of presence in terms of its definition and measurement and no consistent correlation was revealed between presence and task performance in this thesis. On going research towards correlating subjective assessments of perceived presence with physiological responses could offer valuable insights. Perceived presence, though, might not be related to measurable characteristics of human responses and more research needs to be conducted towards understanding this often stated desirable, human perception issue. Generally, formally validated statistical frameworks need to be employed towards that goal.

Monitoring participants' navigation patterns indicated significantly higher amount of idle time for the HMD stereo head tracked condition compared to the real-world condition. For the HMD mono mouse condition, a significant positive correlation was revealed between idle time and accurate spatial memory recall scores. For the desktop condition, a significant correlation was revealed between idle time and accurate memory recall, confidence and prior probabilities related to the 'know' awareness state. Interestingly, the above correlations were revealed in the conditions including a mouse interface. This range of results offers more information about participants' behaviour during the main experiments.

It is evident that techniques based exclusively on task performance towards assessing the simulation fidelity of a VE implementation are quite limited compared with formal frameworks, such as the memory awareness framework which operates on a cognitive rather than on a task point of view. The simulation of reality could be inadequate when only simulation of task performance in the real world is targeted. This thesis shows that the simulation of a space or task situation through computer graphics and VE technologies resulting in similar task performance as the real world does not necessarily guarantee accurate simulation of the cognitive strategies that humans employ in order to complete the task. How this is important depends on the application in question.

6.2 Implications for Future Research

The experiments of Chapter 5 were formally designed. However, certain improvements could be accomplished by the following actions:

- It would be useful for the participants to be administered a spatial ability test (SAT) independently of the experiment in order to statistically abstract out the possible differences in results due to differences in spatial ability. Randomisation of the groups is a measure against this possibility, however, a spatial ability test could ensure that the groups are 'equal' in terms of the participants' spatial ability level.
- The task employed in the experiments of Chapter 5 was designed to allow for minimal navigation to ensure *control* across conditions. A spatial task could employ more interactivity with the computer graphics world, emphasising the 3D aspects of the space and including them in the task. However, control is essential

between conditions for results to be valid and this becomes harder as more interactivity is introduced.

- The memory awareness questionnaire was not administered to participants before the actual experimental studies took place. It might be useful for participants to train on how to follow the instructions given, before the actual experiment starts.
- Since the ‘remember’, ‘know’, ‘familiar’ and ‘guess’ awareness states employed do not represent episodic and semantic memory but different cognitive processes of recollection, they could be just labelled ‘category 1’, ‘category 2’, ‘category 3’ and ‘category 4’ respectively, in order to avoid any preconceptions associated with the actual words used.
- The device (Minolta Chromameter) used in Chapter 5 to measure luminance and chromaticity of the illuminant and the materials in the scene outputs Y, xy tristimulus values. A spectroradiometer is a more sophisticated (and expensive) device which could be employed, resulting in more accurate readings across the visible spectrum.
- The Head Mounted Displays employed in the preliminary study (Hewlett Packard Laboratories prototype) and in the main studies (Kaiser Pro30) allow for periphery vision. An amount of illumination from the displays was adequate for the real world experimental surroundings to be noticed by certain participants, even if the experimental space was darkened for the main studies and any physical light blocked. It would be useful to run such experiments with fully immersive HMDs and a wider FoV.

Several avenues of future work have been highlighted through the course of this thesis:

Computational models: A computational model based on experiments involving human participants could provide an automatic means of assessing simulation fidelity of a VE application without the need for running time consuming experiments with human participants. This does not mean that relevant experimentation should be substituted, but rather such experiments should provide the means of exploiting the human perceptual abilities towards improving computer graphics rendering and VE technologies. Generally, it would be useful to have a scientifically formed framework for feeding results acquired by human experimentation back to the system or algorithm in question.

Comparison between different rendering algorithms: In the preliminary study of this thesis a flat shaded rendering was employed, however, in the main study a photorealistic rendering was used based on photometric data acquired in the real world. Varied quality of rendering was not, however, incorporated into the experimental frameworks employed. It would be of interest, therefore, to explore the variations of cognitive strategies for spatial memory related to different rendering algorithms.

Presence as a correlate of eye movements: The notion of presence is often assessed by means of questionnaires. Although the use of these questionnaires has been validated though years of experimentation, it would be valuable to validate their outcomes to physiological responses. In an application incorporating any kind of task, Areas Of Interest could be assigned as circles or rectangles of the same centre but varying size and participants eye fixations in a space could be correlated to relevant Likert scale ratings. This hypothesis and eye movement theory should be explored, employing these techniques to improve the efficiency of the computer graphics rendering. Relevant research is beginning to emerge.

The study of aspects of human perception in order to exploit them for improving computer graphics rendering and Virtual Environment technologies (displays and interaction devices) is essential for the further development of the computer graphics field. Such research has already resulted in perceptually based algorithms and image quality metrics already as discussed in Chapter 2, however, research as such has not been developed to its full potential. This is a major challenge for computer graphics and will guide the community to exciting avenues by incorporating frameworks and theories of human perceptual mechanisms that have been validated through existing experimentation. This thesis demonstrated how the incorporation of such a framework uncovers aspects of Virtual Environment technology and displays not otherwise apparent and is a step in that direction.

Bibliography

[3ds2001] 3D Studio MAX modelling suite. (2001).
http://www2.discreet.com/index-nf.html

[AB93] Arthur, K.W. & Booth, K.S. (1993). Evaluating 3D Task Performance in Fish Tank Virtual Worlds. *ACM Transactions on Information Systems*, 11(3), 239-265.

[AHC97] Arthur, E.J., Hancock, P.A., Chrysler, S.T. (1997). The Perception of Spatial Layout in Real and Virtual Worlds. *Ergonomics*, 40(1), 69-77.

[AM90] Alfano, P.L., Michel, G.F. (1990). Restricting and Field of View: Perceptual and Performance Effects. *Perceptual and Motor Skills*, 70, 35-45.

[Ari] Aristotle. (384-322 BC). Poetics.

[Art2000] Arthur, K.W. (2000). Effects of FOV on Performance with Head Mounted Displays. *Ph.D thesis*, UNC, USA.

[Au2001] Autocad. (2001). http://www3.autodesk.com/

[Bad97] Baddeley, A. (1997). *Human Memory, Theory and Practice*. Psychology Press.

[BB99] Bystrom, K., Barfield, W. (2000) Collaborative Task Performance for Learning Using a Virtual Environment. *Presence: Teleoperators and Virtual Environments, 8(4),* 435-448. MIT Press.

[BBH99] Bystrom, K-E., Barfiled, W. Hendrix, C. (1999). A Conceptual Model of the Sense of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 8(2), 241-244. MIT Press.

[BE92] Bulthoff, H.H & Edelman, S. (1992). Psychophysical support for a two-dimensional view interpolation theory of object recognition. *Proc. of the National Academy of Sciences of the USA*, 89, 60-64.

[BF95] Barfield, W. & Furness, T., (1995). Introduction to Virtual Environments and Advanced Interface Design. In Barfield, W. & Furness, T., editors, *Virtual Environments and Advanced Interface Design*, Oxford University Press.

[BGG96] Bruce, V., Green, P. R., Georgeson, M.A. (1996). *Visual Perception: Physiology, Psychology and Ecology*. Psychology Press.

[BH95] Barfield, W., Hendrix, C. (1995). The Effect of Update Rate on the Sense of Presence within Virtual Environments. *Journal of the Virtual Reality Society*, 1(1), 3-15.

[BHAW99] Bowman, A., Hodges, L.F., Allison, D., Wineman, J. (1999). The Educational Value of an Information-Rich Virtual Environment. *Presence: Teleoperators and Virtual Environments, 8(3)*, 317-331. MIT Press.

[BHB97] Barfield, W., Hendrix, C., Bystrom, K. (1997) Visualizing the Structure of Virtual Objects Using Head-tracked Stereoscopic Displays. *Proc. of the Virtual Reality Annual International Symposium,* 114-119.

[Bie87b] Biederman, I. (1987b). Matching Image Edges to Object Memory. *Proc. of the First International conference on Computer Vision.* IEEE Computer Society, London.

[BM98b] Bolin, M.R., Meyer, G.W. A (1998). Perceptually based Adaptive Sampling Algorithm. *Proc. of ACM SIGGRAPH 1998*, 299-310.

[Br99] Brooks, F.P. Jr. (1999). What's Real About Virtual Reality. *IEEE Computer Graphics and Applications,* Nov.-Dec. 1999, 16-27.

[BT81] Brewer, W.F. & Treyens, J.C. (1981). Role of Schemata in Memory for Places. *Cognitive Psychology*, 13, 207-230.

[BTG97] Bliss, J.P., Tidwell, P.D., Guest, M.A. (1997). The Effectiveness of Virtual Reality for Administering Spatial Navigation Training to Firefighters. *Presence: Teleoperators and Virtual Environments*, 6(1), 73-86. MIT Press.

[BW93] Barfield, W. and Weghorst, S. (1993). The Sense of Presence within Virtual Environments: A Conceptual Framework. In G. Salvendy and M. Smith (Eds.), *Human-Computer Interaction: Software and Hardware Interfaces*, 699-704. Amsterdam, Netherlands: Elsevier Science Publishers, Orlando, FL.

[BW94] Bailey, J.H. and Witmer, B.G. (1994). Learning and transfer of spatial knowledge in a virtual environment. *Proc. of the Human Factors & Ergonomics Society 38th Annual Meeting,* 1158-1162, Santa Monica, CA: Human Factors & Ergonomics Society.

[BW95] Billinghurst, M. & Weghorst, S. (1995). The Use of Sketch Maps to Measure Cognitive Maps of Virtual Environments. *Proc. of Virtual Reality Annual International Symposium (VRAIS 1995),* 40-47.

[BW2001] Bowman, D.A. & Wingrave, C. (2001). Design and Evaluation of Menu Systems for Immersive Virtual Environments. *Proc. of the IEEE Virtual Reality 2001 conference*, 149-156.

[BZSS95] Barfield, W., Zeltzer, D., Sheridan, T., Slater, M. (1995) Presence and Task Performance within Virtual Environments. In W. Barfield and T.A. Furness III (Eds.), *Virtual Environments and Advanced Interface Design,* 473-513. Oxford University Press.

[Car2000] Cardullo, F.(ed.) (2000). *Flight and Ground Vehicle Simulation Update 2001: course notes*, Binghamton, New York: Watson School of Continuing Education, SUNY Binghamton.

[CG85] Cohen, M.F., Greenberg, D.P. (1985). The Hemi-cube: A Radiosity for Complex Environments. *Proc. of ACM SIGGRAPH 1985,* 19, 31-40.

[CGPAC97] Conway, M.A., Gardiner, J.M., Perfect, T.J., Anderson, S.J., Cohen, G.M. Changes in Memory Awareness During Learning: The Acquisition of Knowledge by Psychology Undergraduates. *Journal of Experimental Psychology,* 126(4), 393-413.

[CMTDM2000] Chalmers, A., McNamara, A., Troscianko, T., Daly, S., Myszkowski, K. (2000). *Image Quality Metrics*, ACM SIGGRAPH 2000, course notes.

[CN98] Cobb, S.V. & Nichols, S. (1998). Static Posture Tests for the Assessment of Postural Instability after Virtual Environment Use. *Brain Research Bullletin*, 47(5), 459-464.

[CNRW99] Cobb, S.V., Nichols, S., Ramsey, A., Wilson, J.R. (1999). Virtual Reality-Induced Symptoms and Effects. *Presence: Teleoperators and Virtual Environments*, 8(2), 169-186. MIT Press.

[Coo95] Cooper, L.A. (1995). Varieties of Visual Representations: How are we to analyse the concept of mental image? *Neuropsychologia*, Vol. 33(11), 1575-1582 Elsevier Science Ltd.

[Coo99] Coolican, H. (1999). *Research Methods and Statistics in Psychology*, 3rd edition. Hodder & Stoughton.

[CSDKH92] Cruz-Neira, C., Sandin, D.J., DeFanti, T.A., Kenyon, R.V., Hart, J.C. (1992). The CAVE: Audio Visual Experience Automatic Virtual Environment, *Communications of the ACM*, 35, 65-72.

[Dal93] Daly, S. (1993). The Visible Differences Predictor: An algorithm for the assessment of image fidelity. In A. Watson, editor, *Digital Image and Human Vision*, 179-206. Cambridge, MA, MIT Press.

[DB98] Darken, R.P. & Banker, W. P. (1998). Navigating in Natural Environments: A Virtual Environment Training Transfer Study. *Proc. of VRAIS 1998*, 12-19.

[DFAB93] Dix, A., Finlay, J., Abowd, G. & Beale, R. (1993). *Human Computer Interaction*. New York: Prentice Hall.

[DKU98] Draper, J.V., Kaber, D.B., Usher, J.M. (1998). Telepresence. *Human Factors: The Journal of the Human Factors Society,* 40(3), 354-375.

[DL97] Dizio, P. & Lackner, J.R. (1997). Circumventing Side Effects of Immersive Virtual Environments. In M. Smith, G. Salvendy, and R. Koubek (Eds.*), Design of computing systems: Social and ergonomic considerations*, 893-896. Amsterdam, Netherlands: Elsevier Science Publishers, San Francisco, CA.

[DM2001] Drago & Myszkowski, K. (2001). Validation proposal for global illumination and rendering techniques. *Computers & Graphics*, 25(3), 511-518.

[DWH99] Dihn, H.Q., Walker, N., Hodges, L.F. (1999). Evaluating the Importance of Multi-Sensory Input on Memory and the Sense of Presence in Virtual Environments. *Proc. of IEEE VR 1999*, 222-228.

[EB94] Ellis, S. R. & Bucher, U. J. (1994). Distance perception of stereoscopically presented virtual objects optically superimposed on physical objects by a head-mounted see-though display. *Proc. Of the Human Factors Society 38th Annual Meeting*, 1300-1304. Santa Monica, CA: Human Factors Society.

[El94] Ellis, S. What are Virtual Environments? (1994). *IEEE Computer Graphics and Applications,* 17-22.

[El95] Ellis, S. (1995). Human Engineering in Virtual Environments. *Proc. of Virtual Reality World 1995*, 295-301.

[El95] Ellis, S. (ed.) (1995). *Pictorial Communication in Virtual and Real Environments*. Taylor and Francis.

[El96] Ellis, S. (1996). Presence of Mind: A Reaction to Thomas Sheridan's 'Further Musings on the Psychophysics of Presence'. *Presence: Teleoperators and Virtual Environments,* 5(2), 247-259. MIT Press.

[El2001] Ellis, S. (2001) The Design of the User-Interface to a Virtual Environment. In Mania, K. (ed.) course notes of *IEEE VR 2001 conference course: 'Usability Evaluation Techniques for Virtual Reality Technologies',* 28-41.

[FAPI99] Freeman, J., Avons, S.E., Pearson, D.E., Ijsselsteijn, W.A.. (1999). Effects of Sensory Information and Prior Experience on Direct Subjective Ratings of Presence. *Presence: Teleoperators and Virtual Environments,* 8(1), 8-13. MIT Press.

[FDFH90] Foley, J. Van Dam. A., Feiner, S., Hughes, J. (1990). *Computer Graphics: Principles and Practice*. Addison-Wesley Publishing Co. ISBN 0-201-12110-7.

[Fly75] Flynn, J.E. (1975). A Study of Lighting as a System of Spatial Cues. *In EDRA-6 workshop on the Psychological Potential of Illumination*, University of Kansas.

[Fly77] Flynn, J.E. (1977). A Study of Subjective Responses to Low Energy and Non-uniform Lighting Systems. *Lighting design and Applications*, 6-15.

[Fur86] Furness, T. A. (1986) The Supercockpit and its Human Factors Challenges. *Proc. Of the 30th Annual Meeting of the Human Factors Society*, Dayton, OH, 48-52.

[Gab97] Gabbard, J.L. (1997). *A Taxonomy of Usability Characteristics in Virtual Envronments*. M.Sc. thesis, Virginia Polytechnic Institute and State University, USA.

[Gar2001] Gardiner, J.M. (2000). Remembering and Knowing. In the E. Tulving and F.I.M. Craik (Eds.) *Oxford Handbook on Memory*. Oxford University Press.

[GG94] Gregg, V.H & Gardiner, J.M. (1994). Recognition Memory and Awareness: A Large Effect of Study-Test Modalities on 'Know' Responses Following a Highly Perceptual Orienting Task. *European Journal of Cognitive Psychology*, 6(2), 131-147.

[GHS99] Gabbard, J.L., Hix, D. Swan II, E. J. (1999). User-Centered Design and Evaluation of Virtual Environments. *IEEE Computer Graphics and Applications*, 51-59.

[Gib71] Gibson, J. (1971). The Information Available in Pictures. *In Leonardo*, vol.4, 27-35. Great Britain: Pergamon Press.

[Gla95] Glassner A.S. (1995). *Principles of Digital Image Synthesis.* Morgan Kaufmann Publishers, San Fransisco, California.

[GM98] Grant, S.C. & Magee, L.E. (1998). Contributions of Proprioception to Navigation in Virtual Environments. *Human Factors: The Journal of the Human Factors Society*, 40(3), 489-497.

[GMN95] Goldenberg, G., Mullbacher, W., Nowak, A. (1995). Imagery without Perception – A Case Study of Anosognosia for Cortical Blindness. *Neuropsychologia*, 33(11), 1373. Elsevier Science Ltd.

[Gr99] Greenberg, D.P. (1999). A Framework for Realistic Image Synthesis. *Communications of the ACM*, 42(8), 45-53.

[GSRH99] Gabbard, J.L., Swartz, K., Richey, K., Hix, D. (1999). Usability Evaluation Techniques: A Novel Method for Assessing the Usability of an Immersive Medical Visualisation VE. *Proc. of VWSIM 1999*, 165-170.

[GW95] Gilkey, R.H. & Weisenberger, J.M. (1995). The Sense of Presence for the Suddenly Deafened Adult: Implications for Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 4(4), 357-363. MIT Press.

[HB96a] Hendrix, C. & Barfield, W. (1996). Presence within Virtual Environments as a Function of Visual Display Parameters. *Presence: Teleoperators and Virtual Environments,* 5(3), 274-289. MIT Press.

[HB96b] Hendrix, C. & Barfield, W. (1996). The Sense of Presence within Auditory Virtual Environments. *Presence: Teleoperators and Virtual Environments,* 5(3), 290-301. MIT Press.

[HBKDN90] Hettinger, L.J., Berbaum, K.S., Kennedy, R.S., Dunlap, W.P., Nolan, M.D. (1990). Vection and Simulator Sickness. *Military Psychology Journal*, 2(3), 171-181.

[HD92] Held, R. & Durlach, N. Telepresence, Time Delay and Adaptation. *NASA Conference publication 10032*, Ch.28, 1992.

[Hee92] Heeter, C. (1992). Being there: The Subjective Experience of Presence. *Presence: Teleoperators and Virtual Environments*, 1(2), 262-271. MIT Press.

[Hen92] Henry, D. (1992). *Spatial Perception in Virtual Environments*. M.Sc. thesis, HITLab, University of Washington, USA.

[HPWG98] Hoffman, H.G., Prothero, J., Wells, M.J., Groen, J. (1998). Virtual Chess: Meaning Enhances User's Sense of Presence in Virtual Environments. *International Journal of Human Computer Interaction*, 10(3), 251-263.

[HS98] Howe, T. & Sharkey, P.M. (1998). Identifying Likely Successful Users of Virtual Reality Systems. *Presence: Teleoperators and Virtual Environments*, 7(3), 308-316.

[Kaj86] Kajiya, J.T. (1986). The Rendering Equation. *In D.C. Evans and R. J. Athay editors, Computer Graphics (Proc. of ACM SIGGRAPH 1986),* 20(4), 143-150.

[Kaj90] Kajiya, J.T. (1990). Radiometry and Photometry for Computer Graphics. *Proc. of ACM SIGGRAPH 1990 Advanced Topics in Ray Tracing course notes.*

[KBDH96] Kennedy, R. S., Berbaum, K. S., Dunlap, W. P., and Hettinger, L. J. (1996). Developing Automated Methods to Quantify the Visual Stimulus for Cybersickness. *Proc. of the Human Factors and Ergonomics Society 40th Annual Meeting*, 1126-1130. Santa Monica, CA: Human Factors & Ergonomics Society.

[KBJ95] Kosslyn, S.M., Behrmann, M., Jeannerod, M. (1995) The Cognitive Neuroscience of Mental Imagery. *Neuropsychologia*, 33(11), 1335-1344. Elsevier Science Ltd.

[KG94] Koriat, A., Goldsmith, M. (1994). Memory in Naturalistic and Laboratory Contexts: Distinguishing the accuracy oriented and quantity oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, 123, 297-315.

[KLBL93] Kennedy, R.S., Lane, N.E., Berbaum, K.S., Lilienthal, M.G. (1993). Simulator Sickness Questionnaire: An Enhanced method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3), 203-220.

[KLLBH92] Kennedy, R. S., Lane, N. E., Lilienthal, M. G., Berbaum, K. S., and Hettinger, L. J. (1992). Profile Analysis of Simulator Sickness Symptoms: Application to Virtual Environment Systems. *Presence: Teleoperators and Virtual Environments*, 1(3), 295-301. MIT Press.

[KMN94] Kashigazi, K.; Matsubara, A., Nagamachi, M. (1994). A Feature Detection Mechanism of Design in Kansei Engineering, *Human Interface*, 9(1), 9-16.

[Kru77] Krueger, M.W. (1977). Responsive Environments. *Proc. of NCC*, 375-85.

[Kru85] Krueger, M.W. (1985). VIDEOPLACE – an Artificial Reality. *Proc. of ACM SIGCHI conference*, 35-40.

[KS97] Kennedy, R.S. & Stanney, K.M. (1997) Aftereffects of Virtual Environment Exposure: Psychometric issues. In M. Smith, G. Salvendy, and R. Koubek (Eds.), *Design of computing systems: Social and ergonomic considerations,* 897-900. Amsterdam, Netherlands: Elsevier Science Publishers, San Francisco, CA.

[KSD2000] Kennedy, R.S., Stanney, K.M., Dunlap, W.P. (2000). Duration and Exposure to Virtual Environments: Sickness Curves During and Across Sessions. *Presence: Teleoperators and Virtual Environments*, 9(5), 463-472. MIT Press.

[KSOD97] Kennedy, R. S., Stanney, K. M., Ordy, J. M., and Dunlap, W. P. (1997). Virtual Reality Effects Produced by Head-Mounted Display (HMD) on Human Eye-hand Co-ordination, Postural Equilibrium, and Symptoms of Cybersickness. *Society for Neuroscience Abstracts*, 23, 772.

[LD94] Lackner, J.R. and DiZio, P. (1994). Rapid adaptation to Coriolis Force Perturbations of Arm Trajectory. *Journal of Neurophysiology*, 72, 299-313.

[LD97] Lombard, M., Ditton, T. (1997). At the heart of it All: The concept of Presence. *JCMC* 3(2), http://www.ascusc.org/jcmc/vol3/issue2/lombard.html

[Li2001] Lightscape radiosity renderer. (2001). http://www.lightscape.com/

[LKGBMB94] Lampton, D.R., Knerr, B.W., Goldberg, S.L., Bliss, J.P., Moshell, J.M., Blau, B.S. (1994). The Virtual Environment Performance Assessment Battery (VEBAP): Development and Evaluation. *Presence: Teleoperators and Virtual Environments*, 3(2), 145-157. MIT Press.

[Loo92] Loomis, J.M. (1992). Presence and Distal Attribution: Phenomenology, Determinants and Assessment. *SPIE Vol. 1666: Human Vision, Visual Processing and Digital Display III*, 590-595.

[LSG94] Lischinski, D., Smits, B., Greenberg, D.P. (1994). Bounds and Error Estimates for Radiosity. In A. Glassner, editor, *Proc. of ACM SIGGRAPH 1994*, 67-74.

[LST2001] Linderman, R.W., Sibert, J.L., Templeman, J.N. (2001). The Effect of 3D Widjet Representation and Simulated Surface Constraints on Interaction in Virtual Environments. *Proc. of IEEE VR 2001*, 141-148.

[Lub95] Lubin, J. (1995). A Visual Discrimination Model for Imaging System Design and Development. In Peli E. (ed.), *Vision Models for target detection and recognition*, 245-283. World Scientific.

[Man2001a] Mania, K. (2001). On Simulation Fidelity and Naturalism of Virtual Interfaces (2001). *Proc. of the IEEE VR 2001 workshop on 'The Future of VR and AR Interfaces: Multi-modal, Humanoid, Adaptive and Intelligent',* GMD Publications.

[Man2001b] Mania, K. (2001). Connections between Lighting Impressions and Presence in Real and Virtual Environments: An Experimental Study. *Proc. of Afrigraph 200.* ACM Publications.

[MC98] Mania, K. & Chalmers, A. (1998). A Classification for User Embodiment in Collaborative Virtual Environments. *Proc. of the 4th International Conference on Virtual Systems and Multimedia (VSMM 98)*, 177-182. IOS Press - Ohmsha, Ltd.

[MC99] Mania, K. & Chalmers, A., (1999). Between Real and Unreal: Investigating Presence and Task Performance, A Pilot Study. *Proc. of 2nd International Workshop on Presence*, University of Essex, Colchester, UK. ISBN 9-386-1571-X.

[MC2000] Mania, K. & Chalmers, A. (2000). A User-Oriented Methodology for Evaluating Virtual Environment Applications. *Proc. of 3rd International Workshop on Presence*, Delft, Netherlands. ISBN 9-386-1571-X.

[MC2001] Mania, K. & Chalmers, A. (2001). The Effects of Levels of Immersion on Presence and Memory in Virtual Environments: A Reality Centred Approach. *Cyberpsychology & Behavior*, 4(2), 247-264.

[Mcn2000] McNamara, A. *Comparing Real and Synthetic Scenes using Human Judgements of Lightness*. Ph.D thesis, Dept. of Computer Science, University of Bristol, UK.

[MCTH2000] Mania, K., Chalmers, A., Troscianko, T., Hawkes, R. (2000). Presence and Task Performance: A Reality Centred Approach. Technical Sketch*, Proc. of ACM SIGGRAPH 2000*, 245.

[MCTH2001] Mania, K., Chalmers, A., Troscianko, T., Hawkes, R. (2001). Simulation Fidelity Metrics for Virtual Environments based on Memory Semantics. Technical Sketch, *Proc. of ACM SIGGRAPH 2001*, 258-258.

[MCTG2000] McNamara, A., Chalmers, A., Troscianko, T., Gilchrist, I. (2000). Comparing Real and Synthetic Scenes using Human Judgements of Lightness. *Proc. of EGWR 2000*, 207-219.

[MEBS2001] Mania, K., Ellis, S.R., Billinghurst, M., Steed, A. -all speakers- (2001). In Mania, K. (ed.) *Usability Evaluation Techniques for Virtual Reality Technologies course notes, IEEE Virtual Reality 2001*.

[Mee2000] Meehan, M. (2000). An Objective Surrogate for Presence: Physiological Response. *Proc. of 3rd International Conference on Presence*. ISBN 9-386-1571-X.

[MJSP2000] Mizell, D.W., Jones, S.P., Slater, M., Spanlang, B. (2000) The influence of Immersion in Virtual Environments on Performance of a Modelling Task. *Report for EPSRC grant GR/M85692*, UK.

[ML92] Meyer, G.W. & Liu, A. (1992). Color Spatial Acuity Control of a Screen Subdivision Image Synthesis Algorithm. *Human Vision, Visual Processing and Digital Display,* 1666(3), 387-399.

[MRCG86] Meyer, G.W., Rushmeier, H. E., Cohen, M.F., Greenberg, D.P. (1986). An Experimental Evaluation of Computer Graphics Imagery. *ACM Transactions on Graphics*, 5(1), 30-35.

[MWJR97] Mon-Williams, M., Wann, J.P., Jenkinson, M., and Rushton, K. (1997). Synesthesia in the Normal Limb. *Proc. of the Royal Society Series B.*, 264, 1007-1010.

[MWR95] Mon-Williams, M., Wann, J.P., Rushton, S. (1995). Binocular Vision in Stereoscopic Virtual Reality Systems. *Digest of the Society for Information Display*, 1, 361-363.

[Mys98b] Myszkowski, M. The Visible Differences Predictor: Applications to Global Illumination Problems. (1998). In G. Drettakis and N. Max, editors, *Proc. of the Eurographics Rendering Workshop 1998*, 233-236. NY, Springer Wien.

[Nag2001] Nagamachi, M. (2001). Perspectives of Kansei Engineering/Kansei Ergonomics. *Proc. of the HCI International Conference 2001.*

[Nic99] Nichols, S. (1999). Physical Ergonomics of Virtual Environment Use. *Applied Ergonomics Journal*, 30, 79-90.

[NJE94] Nemire, K., Jacoby, R.H., Ellis, S.R. (1994). Simulation Fidelity of a Virtual Environment Display. *Human Factors: The Journal of the Human Factors Society*, 36(1), 79-93.

[Pro98] Prothero, J.D. (1998). *The Role of Rest Frames in Vection, Presence and Motion Sickness*. Ph.D thesis, HITLab, University of Washington, USA.

[PPW97] Pausch, R., Proffitt, D., Williams, G. (1997). Quantifying Immersion in Virtual Reality. *Proc. of ACM SIGGRAPH 1997,* 13-18.

[PVG95] Peruch, P., Vercher, J.L, Gouthier, G.M. (1995). Acquisition of Spatial Knowledge through Visual Exploration of Simulated Environments. *Ecological Psychology*, 7(1), 1-20.

[Raj96] Rajaram, S. (1996). Perceptual Effects on Remebering: Recollective Processes in Picture Recognition Memory. *Journal of Experimental Psychology*, 22(2), 365-377.

[RCD97] Robertson, R., Czerwinski, M., Dantzich, M. (1997). Immersion in Desktop Virtual Reality. *Proc. of UIS 1997*, 11-19.

[Rie97] Riesberg, D. (1997). *Cognition: Exploring the Science of the Mind.* W.W Norton and Company, London, UK.

[Rin96] Rinalducci, E.J. (1996). Characteristics of Visual Fidelity in the Virtual Environment. *Presence: Teleoperators and Virtual Environments*, 5(3), 330-345. MIT Press.

[RGPSR95] Rushmeier, H., Ward, G., Piatco, C., Sanders, P., Rust, B. (1995). Comparing Real and Synthetic Images: Some ideas about metrics. *Proc. of EGRW 1995*.

[RMCH97] Rinalducci, E.J., Mapes, D., Cinq-Mars, S.G., and Higgins, K.E. (1997). Determining the Field of View in HMDs: A Psychophysical Method. *Presence: Teleoperators and Virtual Environments*, 5(3), 353-356. MIT Press.

[RMW94] Rushton, S. Mon-Williams, M & Wann, J.P. (1994). Binocular Vision in a Bi-ocular world: New generation Head Mounted Displays avoid causing visual deficit. *Displays*, 15(4), 255-260.

[Rog95] Rogers, S. (1995). Perceiving Pictorial Space. In W. Epstein & S. Rogers, editors*, Perception of Space and Motion*. San Diego: Academic Press.

[RPJ99] Ruddle, R. A., Payne, S. J., Jones, D. M. (1999). Navigating Large-Scale Virtual Environments: What Differences Occur Between Helmet-Mounted and Desk-Top Displays?. *Presence: Teleoperators and Virtual Environments, 8(2),* 157-168. MIT Press.

[RS86] Rolfe, J.M. & Staples, K.J. (1986). *Flight Simulation.* London, Cambridge University Press.

[SAMG97] Singer, M.J., Allen, R.C., McDonald, D.P., and Gildea, J.P. (1997). Terrain Appreciation in Virtual Environments: *Spatial knowledge acquisition (Technical Report 1056)*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Science.

[Sch95] Schloerb, D.W. (1995). A Quantitative Measure of Telepresence. *Presence: Teleoperators and Virtual Environments,* 4(1), 64-80. MIT Press.

[SDLS97] Salzman, M. C., Dede, C., Loftin, R. B., Sprague, D. (1997) Assessing Virtual Reality Potential for Teaching Abstract Science. *Proc. of the Human Factors and Ergonomics Society 41st Annual Meeting*, 1208-1212.

[SF70] Segal, S.J. & Fusella, V. (1970). Influence of Imaged Pictures and Sounds on Detection of Visual and Auditory Signals. *Journal of Experimental Psychology*, 83, 458-464.

[SG95] So, R. H. and Griffin, M. J. (1995). Effects of Lags on Human Operator Transfer Functions with Head-Coupled Systems. *Aviation, Space, and Environmental Medicine Journal*, 66(6), 550-556.

[She92] Sheridan, T.B. (1992). Further Musings on the Psychophysics of Presence. *Presence: Teleoperators and Virtual Environments,* 5(2), 241-246. MIT Press.

[Sla99] Slater, M. (1999). Measuring Presence: A Response to the Witmer and Singer Questionnaire. *Presence: Teleoperators and Virtual Environments,* 8(5), 560-565. MIT Press.

[SM71] Shepard, R.N. & Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. *Science*, 171, 701-703.

[Smi87] Smith, R. (1987) Experiences with the Alternate Reality Kit: An Example of the Tension between Literalism and Magic User Interface Metaphore. *Proc. of ACM CHI+GI 1987 on Human Factors in Computing Systems and Graphics*, *Interface*, 61-67.

[Smi92] Smith, M.E. (1992). Neurophysiological Manifestations of Recollective Memory Experience during Recognition Memory Judgements. *Journal of Cognitive Neuroscience*, 5, 1-13.

[SMK98] Stanney, K.M, Mourant, R.R., Kennedy, R.S. (1998). Human Factors Issues in Virtual Environments: A Review of the Literature. *Presence: Teleoperators and Virtual Environments*, 7(4), 327-351, MIT Press.

[Sno96] Snow, M.P. (1996*). Charting Presence in Virtual Environments and Its Effects on Performance*. Ph.D thesis in Industrial and Systems Engineering, Virginia Tech, USA.

[SS94] Stanney, K.M & Salvendy, G. (1994). Effects of Diversity in Cognitive Restructuring Skills on Human-Computer Performance. *Ergonomics*, 37(4), 595-609.

[SS2000] Slater, M. & Steed, A. (2000). A Virtual Presence Counter. *Presence: Teleoperators and Virtual Environments*, 9(5), 413-434. MIT Press.

[SSetal98] Stanney, K.M., Salvendy, G., Deisigner, J., DiZio, P., Ellis, S., Ellison, E., Fogleman, G., Gallimore, J., Hettinger, L., Kennedy, R., Lackner, J., Lawson, B., Maida, J., Mead A., Mon-Williams, M., Newman, D., Piantanida, T., Reeves, L., Riedel, O., Singer, M., Stoffregen, T., Wann, J., Welch, R., Wilson, J., Witmer, B. (1998). Aftereffects and Sense of Presence in Virtual Environments: Formulation of a research and development agenda. Report sponsored by the Life Sciences Division at NASA Headquarters. *International Journal of Human-Computer Interaction*, 10(2), 135-187.

[SSK2001] Sato, Y., Saito, M., Koike, H. (2001). Real-Time Input of 3D Pose and Gestures of a User's Hand and its Applications for HCI. *Proc. of IEEE VR 2001*, 79-86.

[SSMM98] Slater, M., Steed, A., McCarthy, J., Maringelli, F. (1998). The Influence of Body Movement on Subjective Presence in Virtual Environments. *Human Factors: Journal of the Human Factors Society,* 40(3), 469-477.

[SSUS2000] Slater, M., Sadagic, A., Usoh, M., Schroeder, R. (2000). Small Group Behavior in a Virtual and Real Environment, A Comparative Study. *Presence: Teleoperators and Virtual Environments*, 9(1), 37-51. MIT Press.

[SU94] Slater, M. & Usoh M. (1994). Depth of Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments,* 3(2), 130-144. MIT Press.

[SULK96] Slater, M., Usoh, M., Linakis, V., Kooper, R. (1996). Immersion, Presence and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess. *Proc. of VRST 1996*, 163-171.

[SUS95] Slater, M., Usoh, M., Steed, A. (1995). Taking Steps: The influence of a Walking Technique on Presence in Virtual Reality. *ACM Transactions on Computer Human Interaction*, 2, 201-219.

[SW99] Singer, M.J., Witmer, B.G. (1999). On Selecting the Right Yardstick. *Presence: Teleoperators and Virtual Environments,* 8(5), 566-573. MIT Press.

[SW95] Slater, M. & Wilbur, S. (1995). Through the Looking Glass world of Presence: A Framework for Immersive Virtual Environments. In M. Slater (ed.), *Proc. of FIVE 1995 Framework for Immersive Virtual Environments*. QMW University of London.

[SW98] Snow, M.P., Williges, R.C. (1998). Empirical Models Based on Free-Modulus Magnitude Estimation of Perceived Presence in Virtual Environments. *Human Factors: Journal of the Human Factors Society,* Vol. 40, No. 3, 386-402.

[Su65] Sutherland, I.E. (1965). The Ultimate Display. *International Federation of Information Processing*, vol. 2, 506-508.

[Su70] Sutherland, I.E. (1970). Computer Displays. *Sci. American*, vol. 222, 56-81.

[Tal2001] Tallyn, E. (2001). *Applying Narrative Theory to the Design of Digital Interactive Narratives with the Aim of Understanding and Enhancing Participant Engagement.* Ph.D thesis, University of the West of England, UK.

[TMPV98] Thornton, W.E., Moore, T.P., Pool, S.L. and Vanderploeg, J. (1998). Clinical Characterization and Etiology of Space Motion Sickness. *Aviation Space & Environmental Medicine,* 58(Suppl.), A1-A8.

[TR93] Tumblin, J. & Rushemeier, H.E. (1993). Tone Reproduction for Realistic Images. *IEEE Computer Graphics and Applications*, 13(6), 42-48.

[Tra91] Travis, D. (1991). *Effective Color Displays*. Academic Press.

[Tul85] Tulving, E. (1985). Memory and Concioussness. *Canadian Psychologist, 26*, 1-12.

[Tul93] Tulving, E. (1993). Varieties of Concioussness and Levels of Awareness in Memory. In A.D. Baddeley and L. Weiskrantz (Eds.), *Attention: Selection, Awareness and Control.* A tribute to Donald Broadbent, 283-299. London: Oxford University Press.

[UAWBSSB99] Usoh, M., Arthur, K., Whitton, M.C., Bastos, R., Steed, A., Slater, M., Brooks, F.P. (1999). Walking>Walking-in-Place>Flying in Virtual Environments. *Proc. of ACM SIGGRAPH 1999*, 359-364.

[UCAS2000] Usoh, M., Catena, E., Arman, S., Slater, M. (2000) Using Presence Questionnaires in Reality. *Presence: Teleoperators and Virtual Environments*, 9(5), 497-503. MIT Press.

[VRML97] Virtual Reality Modelling Language 97, http://www.web3d.org/technicalinfo/specifications/vrml97/index.htm

[WBLMS96] Welch, R.B., Blackmon, A., Lin, A., Mellers, B.A., Stark, L.W. (1996). The Effects of Pictorial Realism, Delay of Visual Feedback and Observer Interactivity on the Subjective Sense of Presence. *Presence: Teleoperators and Virtual Environments*, 5(3), 263-273. MIT Press.

[WBK96] Witmer, B.G., Bailey, J. H., Knerr, B.W. (1996). Virtual Spaces and Real World places: Transfer of Route Knowledge. *Human-Computer Studies*, 45, 413-428.

[Wel97] Welch, R. (1997). The Presence of Aftereffects In M. Smith, G. Salvendy, and R. Koubek (Eds.), *Design of computing systems: Social and ergonomic considerations,* 273-276. Amsterdam, Netherlands: Elsevier Science Publishers, San Francisco, CA. ISBN: 044482183X.

[Wel99] Welch, R.B. (1999). How Can We Determine if the Sense of Presence Affects Task Performance? *Presence: Teleoperators and Virtual Environments*, 8(5), 574-577. MIT Press.

[WHK98] Waller, D., Hunt, E., Knapp, D. (1998). The Transfer of Spatial Knowledge in Virtual Environment Training. *Presence: Teleoperators and Virtual Environments,* 7(2), MIT Press.

[WM96] Wann, J. & Mon-Williams, M. (1996). What does Virtual Reality NEED? Human Factors Issues in the Design of Three-Dimensional Computer Environments. *International Journal of Human-Computer Studies*, 44, 829-847.

[WNH97] Wilson, J.R., Nichols, S., and Haldane, C. (1997). Presence and Side Effects: Complementary or Contradictory? In M. Smith, G. Salvendy, and R. Koubek (Eds.), *Design of computing systems: Social and ergonomic considerations*, 889-892. Amsterdam, Netherlands: Elsevier Science Publishers, San Francisco, CA. ISBN: 044482183X.

[Wor2001] World UP software. (2001). http://www.sense8.com/

[WRM95] Wann, J.P. Rushton, S.K., and Mon-Williams M. (1995). Natural Problems for Stereoscopic Depth Perception in Virtual Environments. *Vision Research*, 19, 2731-2736.

[WS94] Witmer, B.G. and Singer, M.J. (1994). Measuring Immersion in Virtual Environments *(Technical Report 1014).* Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Science.

[WS98] Witmer, B. G., Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments,* 7(3), 225-240. MIT Press.

[WS98] Witmer. B. G., Sadowski, W.J. Jr. (1998). Non-Visually Guided Locomotion to a Previously Viewed Target in Real and Virtual Environments. *Human Factors: The Journal of the Human Factors Society,* 40(3), 478-488.

[WWH94] Williams, H. P, Wickens, C.D., & Hutchinson, S. (1994). Fidelity and Interactivity in Navigational Training: A Comparison of Three Methods. *Proc. of the Human Factors Society 38th meeting*, 1163-1167. Santa Monica CA, Human Factors Society.

[Zel92] Zeltzer, D. (1992). Autonomy, Interaction and Presence. *Presence: Teleoperators and Virtual Environments,* 1(1), 127-132. MIT Press.

[ZJ98] Zahoric, P & Jenison, R.L. (1998). Presence and Being-in-the-World. *Presence: Teleoperators and Virtual Environments*, 7(3), 225-240. MIT Press.

Appendix A

Materials

A1. Materials' Pack for the Preliminary Study (Chapter 4)

QUESTIONNAIRE 1

INSTRUCTIONS

In this questionnaire there are 22 questions which are of the following form:

1 *Where were the Boers from?*

❒ England ❒ France ❒ South Africa ❒ Italy

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

Please, select **one** answer and put a cross in the box next to that answer. (see above)

Next indicate how confident you are that your answer is CORRECT by placing a cross in one of the confident boxes. (see above). Finally indicate the basis for your answer.

It is very important that you respond accurately. You may have selected an answer because:
1) You remembered a specific episode from the seminar. In this case you might have images and feelings in mind relating to the recalled information. Perhaps you virtually 'hear' again

or ‘see’ again the lecturer presenting some item of information or remember visually the specific slide that information was included into. Answers such as these are called REMEMBER answers.

2) You might just ‘know’ the correct answer and the alternative you have selected just ‘stood out’ from the choices available. In this case you would not recall a specific episode and instead you would simply know the answer. Answers with this basis are called KNOW answers.

3) It may be, however, that you did not remember a specific instance, nor do you know the answer. Nevertheless, the alternative you have selected may seem or feel more familiar than any of the other alternatives. Answers made on this basis are called FAMILIAR answers.

4) You may not have remembered, known, or felt that the choice you selected have been familiar. In which case you may have made a guess, possibly an informed guess, e.g. some of the choices look unlikely for other reasons so you have selected the one that looks least unlikely. This is called a GUESS answer.

Indicate the psychological basis for your answer by checking ONE of the boxes for REMEMBER, KNOW, FAMILIAR or GUESS.

Questions related to the seminar room

1 *What was the colour of the cube in the seminar room on the table?*

❐ Blue ❐ Red ❐ Green ❐ Black

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

2 Where was the conference advertised on the poster on the wall to be held?

❐ Spain ❐ Germany ❐ Portugal ❐ Bristol

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

3 What was the colour of the folder on the table?

❐ Green ❐ Brown ❐ Red ❐ White

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

4 How many separate panes of glass are there in the room?

❒ 7 ❒ 6 ❒ 4 ❒ 12

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

5 How many walls of the seminar room have windows?

❒ 2 ❒ none ❒ 1 ❒ 3

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

6 *Which was the colour of the carpet on the floor in the seminar room?*

❒ Red ❒ Blue ❒ Yellow ❒ Magenta

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

Presence Questionnaire for the Preliminary Study

Slight changes were made in the questionnaire's body taken from [SSMM98] to accommodate for the different conditions and context of the experimental study. For instance, wording as 'images that I saw' were replaced with 'sounds that 'I've heard' for the audio-only condition.

QUESTIONNAIRE 2

INSTRUCTIONS

The following questions relate to your experience. Please, circle the appropriate step on the scale from 1 to 7, for each question. In this questionnaire, there are 15 questions of the following form:

1 Please rate how thirsty you are feeling at this moment

not at all 1 2 3 4 5 6 7 very much

The mark close to 'very much' indicates that you are quite thirsty indeed.

QUESTIONS

1 Please rate the extent to which you were aware of background sounds in the room in which this experience was actually taking place. Rate this on the following scale from 1 to 7 (where for example 1 means that you were hardly aware at all of any background sounds):

not at all 1 2 3 4 5 6 7 very much

2 How dizzy, sick or nauseous did you feel resulting from the experience, if at all? Please, answer on the following 1 to 7 scale.

not at all 1 2 3 4 5 6 7 very much so

3 Gender

❒ Male ❒ Female

4 Please rate your sense of being in the seminar room, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place.

not at all 1 2 3 4 5 6 7 very much

5 To what extent were there times during the experience when the seminar room was the reality for you?

at no time 1 2 3 4 5 6 7 most of the time

6 Status

❒ Undergraduate student ❒ Masters student ❒ PhD student
❒ Research Assistant/Fellow ❒ Staff member/technical staff ❒ Faculty
❒ Other (Please specify)…………………………………………

7 When you think back about your experience, do you think of the seminar room more as images that you saw, or more as somewhere that you visited?

images that I saw 1 2 3 4 5 6 7 a place that I visited

8 Have you ever experienced virtual reality/3D applications/games?

never 1 2 3 4 5 6 7 a great deal

9 During the time of the experience, which was the strongest on the whole, your sense of being in the seminar room or of being elsewhere?

being elsewhere | 1 | 2 | 3 | 4 | 5 | 6 | 7 | being there

10 Overall, how well do you think that you achieved the experimental task?

not very well at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | very well

11 Consider your memory of attending the seminar. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By 'structure of the memory' consider things like the extent to which you have a visual memory of the field, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic your imagination, and other such structural elements.

not all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | very much so

12 To what extent do you use a computer in your daily activities?

not all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | very much so

13 During the time of the experience, did you often think to yourself that you were actually in the seminar room?

not very often | 1 | 2 | 3 | 4 | 5 | 6 | 7 | very often

14 Were you involved in the experimental task to the extent that you lost track of time?

not at all | 1 | 2 | 3 | 4 | 5 | 6 | 7 | very much so

15 Further comments

Please write down any further comments that you wish to make about your experience. What things helped to give you a sense of 'really being' in the space, and what things acted to 'pull you out' of this?

Thank you very much for participating in this study. All the answers will be treated entirely confidentially.

A2. Materials' Pack for the Main Studies (Chapter 5)

QUESTIONNAIRE 1

INSTRUCTIONS

Name:
Series:

This is an example:

1 *Object Location Number:*_____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

Please, **fill in the location number** according to the diagram, select **one** answer and put a cross in the box next to that answer. (see above)

Next **indicate how confident you are** that your answer is CORRECT by placing a cross in one of the confident boxes. (see above).

It is very important that you respond accurately. Your awareness could be:

1) **REMEMBER** means that you can visualise clearly the object in the room in your head, in that particular location. You virtually 'see' again elements of the room in your mind.

2) **KNOW** means that you just 'know' the correct answer and the alternative you have selected just 'stood out' from the choices available. In this case you can't visualise the specific image or information in your mind.

3) **FAMILIAR** means that you did not remember a specific instance, nor do you know the answer. It may seem or feel more familiar than any of the other alternatives.

4) **GUESS** means that you may not have remembered, known, or felt that the choice you selected have been familiar. You may have made a guess, possibly an informed guess, e.g. you have selected the one that looks least unlikely.

Please check ONE of the boxes for REMEMBER, KNOW, FAMILIAR or GUESS.

QUESTIONS

1 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

2 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

3 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

4 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

5 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

6 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

7 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

8 *Object Location Number:*____ ❐ Box ❐ Sphere ❐ Pyramid

Confidence: ❐ No conf. ❐ Low conf. ❐ Moderate conf. ❐ Confident ❐ Certain
Awareness: ❐ Remember ❐ Know ❐ Familiar ❐ Guess

9 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

10 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

11 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

12 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

13 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

14 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

15 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

16 *Object Location Number:*____ ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

17 *Object Location Number:____* ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

18 *Object Location Number:____* ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

19 *Object Location Number:____* ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

20 *Object Location Number:____* ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

21 *Object Location Number:____* ❒ Box ❒ Sphere ❒ Pyramid

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

Confidence: ❒ No conf. ❒ Low conf. ❒ Moderate conf. ❒ Confident ❒ Certain
Awareness: ❒ Remember ❒ Know ❒ Familiar ❒ Guess

QUESTIONNAIRE 2

INSTRUCTIONS

The following questions relate to your experience. Please, circle the appropriate step on the scale from 1 to 7, for each question. In this questionnaire, there are 15 questions of the following form:

1 Please rate how thirsty you are feeling at this moment

not at all 1 2 3 4 5 6 7 very much

The mark close to ‘very much’ indicates that you are quite thirsty indeed.

QUESTIONS

1 Please rate the extent to which you were aware of background sounds in the room in which this experience has actually taken place. Rate this on the following scale from 1 to 7 (where for example 1 means that you were hardly aware at all of any background sounds):

not at all 1 2 3 4 5 6 7 very much

2 How dizzy, sick or nauseous did you feel resulting from the experience, if at all? Please, answer on the following 1 to 7 scale.

not at all 1 2 3 4 5 6 7 very much so

3 Gender

❐ Male ❐ Female

4 Please rate your sense of being in the 3D room, on the following scale from 1 to 7, where 7 represents your normal experience of being in a place.

not at all 1 2 3 4 5 6 7 very much

5 To what extent were there times during the experience when the 3D room was the reality for you?

at no time 1 2 3 4 5 6 7 most of the time

6 Status

❐ Undergraduate student ❐ Masters student ❐ PhD student
❐ Research Assistant/Fellow ❐ Staff member/technical staff ❐ Faculty
❐ Other (Please specify)……………………………………………

7 When you think back about your experience, do you think of the 3D room more as images that you've seen, or more as somewhere that you visited?

images I've seen 1 2 3 4 5 6 7 a place that I visited

8 Have you ever experienced virtual reality/3D applications/games?

never 1 2 3 4 5 6 7 a great deal

9 During the time of the experience, which was the strongest on the whole, your sense of being in the 3D room or of being elsewhere?

being elsewhere 1 2 3 4 5 6 7 being there

10 Overall, how well do you think that you achieved the experimental task?

not very well at all 1 2 3 4 5 6 7 very well

11 Consider your memory of the 3D room. How similar in terms of the structure of the memory is this to the structure of the memory of other places you have been today? By 'structure of the memory' consider things like the extent to which you have a visual memory of the field, whether that memory is in colour, the extent to which the memory seems vivid or realistic, its size, location in your imagination, the extent to which it is panoramic your imagination, and other such structural elements.

not all 1 2 3 4 5 6 7 very much so

12 To what extent do you use a computer in your daily activities?

not all 1 2 3 4 5 6 7 very much so

13 During the time of the experience, did you often think to yourself that you were actually in the 3D room?

not very often 1 2 3 4 5 6 7 very often

14 Were you involved in the room experimental task to the extent that you lost track of time?

not at all 1 2 3 4 5 6 7 very much so

QUESTIONNAIRE 3

INSTRUCTIONS

Please, circle the appropriate term, to describe your physical state, for each symptom.

General Discomfort:	none	slight	moderate	severe
Fatigue:	none	slight	moderate	severe
Headache:	none	slight	moderate	severe
Eyestrain:	none	slight	moderate	severe
Difficulty Focusing:	none	slight	moderate	severe
Increased Salivation:	none	slight	moderate	severe
Sweating:	none	slight	moderate	severe
Nausea:	none	slight	moderate	severe
Difficulty Concentrating:	none	slight	moderate	severe
Fullness of Head:	none	slight	moderate	severe
Blurred Vision:	none	slight	moderate	severe
Dizzy (Eyes open):	none	slight	moderate	severe
Dizzy (Eyes closed):	none	slight	moderate	severe
Vertigo:	none	slight	moderate	severe
Stomach Awareness:	none	slight	moderate	severe
Burping:	none	slight	moderate	severe

QUESTIONNAIRE 4

The following questions relate to your impression of the 3D room. Please, circle the appropriate step on the scale from 1 to 7, for each question.

spacious	1	2	3	4	5	6	7	confined
relaxing	1	2	3	4	5	6	7	tense
bright	1	2	3	4	5	6	7	dim
stimulating	1	2	3	4	5	6	7	subduing
dramatic	1	2	3	4	5	6	7	diffuse
uniform	1	2	3	4	5	6	7	non-uniform
interesting	1	2	3	4	5	6	7	uninteresting
radiant	1	2	3	4	5	6	7	gloomy
large	1	2	3	4	5	6	7	small
like	1	2	3	4	5	6	7	dislike
simple	1	2	3	4	5	6	7	complex
uncluttered	1	2	3	4	5	6	7	cluttered
warm	1	2	3	4	5	6	7	uncomfortable
pleasant	1	2	3	4	5	6	7	unpleasant
comfortable	1	2	3	4	5	6	7	cold

Further comments

Please write down any further general comments that you wish to make about your experience. For example, what things helped to give you a sense of 'really being' in the 3D room, and what things acted to 'pull you out' of this?

Thank you very much for participating in this study. All the answers will be treated entirely confidentially.