The Effects of Fidelity on Navigation in Virtual Environments

by

Simon Lessels

Submitted in accordance with the requirements
for the degree of Doctor of Philosophy.

The University of Leeds
School of Computing

July 2005

Abstract

Virtual environments (VEs) offer huge potential for a wide range of applications including the transfer of spatial knowledge from virtual spaces to real world places; beneficial in situations where it would be impractical, too expensive or dangerous, to acquire that knowledge from the real environment. Research has shown that people can acquire near perfect spatial knowledge about real world environments from three-dimensional (3D) VEs. However, the rate of learning is substantially slower, and the information less accurate, than that acquired from the real world. It is often assumed that poor navigational ability in VEs is due to the reduced *fidelity* of the VE system, fidelity is defined as how closely the various components of the VE system resemble those of the real world. This thesis attempts to better understand the effects of, and the relationship between, three aspects of VE fidelity, field of view, visual scene characteristics and the movement interface. Four experimental studies showed that a wide FOV, a high fidelity visual scene, and a simple movement interface, modestly increased participants' ability to navigate efficiently in a desktop VE. However, a study that required participants to physically walk around a VE, displayed via a tracked head mounted display (HMD), showed dramatic performance benefits over the use of stationary desktop displays, and a rotationally tracked HMD that required abstract input for translational movement. Proprioceptive and vestibular feedback allowed participants to navigate a VE as efficiently as they did in a real world study. The potential of VEs for spatial applications, such as learning real world spaces, will not be realised without understanding the effects of the VE system on participants' performance and behaviour. The studies reported in this thesis not only provide much needed empirical results that could be of great benefit to VE application designers, but will also be of interest to researchers investigating human navigation.

Dedication

To my dad.

19th May 1935 to 9th July 2005

Acknowledgements

Firstly, I would like to thank my supervisor Dr Roy A. Ruddle for his intellectual input, support, and patience. I could not have written this thesis without his encouragement, professionalism, and guidance. Thanks also to Professor Peter Dew and Professor Ken Brodlie for their encouragement and guidance.

Thanks to the EPSRC for funding the research. Thanks to Russell Murray for his generous gift of my first PC, many years ago, that started this whole journey. Also, thanks to my family for their generous support over the years.

I would like to thank the various members of the Biosystems group who have made my time at Leeds so enjoyable, Chris Buckley, Dr. Seth Bullock, Dr. John Cartlidge, Dave Harris, Dr. Jason Noble, and Hywell Williams. Thanks to Royce Neagle and all the other members of the Informatics Network for making the research environment so rewarding. Thanks to John Hodrien and Martin Thompson for their technical assistance in the experiments.

Finally, I would like to thank my friends James Bell, Robbie Brodie, and Paul McAusland, also, students and staff of Leeds University, who participated in my experiments, it was a lot of fun, thanks.

Declarations

Some parts of the work presented in this thesis have been published in the following articles:

Ruddle, R A; Lessels, S., Three levels of metric for evaluating wayfinding in: *Proceedings of Virtual Reality Design and Evaluation Workshop.* 2004. Available on CD (ISBN 0 85358 123 1).

Lessels, S; Ruddle, R A., Changes in navigational behaviour produced by a wide field of view and a high fidelity visual scene in: *Proceedings of the 10th Eurographics Symposium on Virtual Environments (EGVE'04)*, pp. 71-78. Eurographics Association. 2004.

Lessels, S; Ruddle, R A., Movement around real and virtual cluttered environments. *Presence: Teleoperators and Virtual Environments.* The MIT Press. In Press.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Research has shown that people can acquire spatial knowledge about real world environments from three-dimensional (3D) virtual environments (VEs). However, the rate of learning is substantially slower, and the information less accurate, than that acquired from the real world. The overall goal of this thesis is to study why people have difficulty navigating in VEs.

The investigative scope of this thesis is limited to the effects on participants' performance and behaviour from three aspects of the VE system, namely: field of view (FOV), visual fidelity, and the movement interface. Four experimental studies showed that a wide FOV, a high fidelity visual scene, and a simple movement interface, modestly increased participants' ability to navigate efficiently in a desktop VE. However, a study that required participants to physically walk around a VE, displayed via a tracked head mounted display (HMD), showed dramatic performance benefits over the use of stationary desktop displays, and a rotationally tracked HMD that required abstract input for translational movement.

This introductory chapter begins with a discussion on VEs to highlight the importance of VE navigation; a brief examination of existing uses of VEs in industry is then presented. The chapter then summarises VE technology that is relevant to navigation and looks at various existing input device taxonomies. By focusing on the most crucial dimensions, a new taxonomy is created that directly relates to navigation within virtual environments. An outline of the various input devices available for navigation is presented, leading into a discussion on how software can interpret the data from the input device to create different interaction techniques for navigation.

The focus then moves to VE display devices; a new display taxonomy is created, within which specific devices are located and discussed. These two taxonomies, covering input and output, capture the design space for VE devices, and show which other devices can be used for similar tasks. Fundamental differences in the characteristics of input and output devices can be shown by their position within their respective taxonomy, differences that are difficult to communicate through a written description, and which are shown in this thesis to affect participants' ability to perform a spatial task in a VE. Chapter One then concludes with brief comments about the changes in rendering technology that have occurred over the last 10 years; advances in computer graphics have allowed visually more complex scenes to be displayed which could potentially affect participants' perception of the VE and so influence task performance.

1.1 Virtual Environments and Navigation

Virtual environments offer huge potential for a wide range of types of application, in particular VEs may be used to:

1. Train for real world tasks, specially where those situations are rare, remote, or dangerous (Waller, Hunt & Knapp, 1998); see table 1.1.

2. Evaluate designs. A VE can be used to examine proposed designs of objects, environments, and situations to gain an understanding of the impact of design decisions, see table 1.2.

3. Visualise data from physical objects (scientific visualisation), or from abstract sources (information visualisation), see table 1.3.

4. Entertain and amuse. In recent years the younger generations have decreased the number of hours watching television in favour of more interactive entertainment media, such PC based computer games and console games, see table 1.4.

In many training applications, a VE is used to simulate the real world, so it is desirable that there is as close a match as possible between the virtual environment and its real world equivalent. Examples range from training surgical skills, to the tactics to be used in a military battle, and routes to be followed during a rescue mission or evacuation. Virtual design applications replicate products and processes that will subsequently be created in a physical form, and an integral part of the design process is the need to move around the scene and view it from a variety of

Table 1.1: VE Applications (training).

Application	**Purpose**
Training	
Hubble space telescope	To train the ground support crew about the Hubble space telescope repair procedures (Loftin & Kenney, 1995).
Helicopter pilots and crew	To train pilots in navigation skills (Sullivan, Darken & McLean, 1998). To train voice marshalling skills for search and rescue operations, Virtalis Ltd (www.virtalis.com)).
Aircraft pilots	Commercial airline simulators, used by many of the large commercial airline companies to train pilots about flight procedures and navigation for over 20 years.
Military training	To train military personnel to operate machinery under various conditions, used by aircraft maintenance crews, tank operators and drivers. VEs have also been used for battle group and infantry training, and close range weaponry tactics and training (Urban Terrain Module, part of the Army's Joint Fires and Effects Trainer System, or JFETS). Also used to train for the detection and disarming of land minds.
Vehicle driving and machine operation	To train drivers of large trucks using scenario based exercises (Bonakdarian, Cremer, Kearney & Willemsen, 1998; Bruzzone, Brandolini & Viazzo, 2004). To train in the procedures for, for example, overhead crane operators (Wilson, Mourant, Li & Xu, 1998).
Ship and submarine piloting	To train ship and submarine piloting teams in harbour and channel ship handling in various geographical and environmental conditions, (VESUB by NAVAIR Orlando).
Firefighters	To train firefighters in the techniques of search and rescue. (Bliss, Tidwell & Guest, 1997).
Air traffic control	To train personnel in ATC procedures, for example the Enhanced Tower Simulator at Fort Rucker, USA.
Surgery skills training	To train laparoscopic surgery skills (Grantcharov, Kristiansen, Bendix, Bardram, Rosenberg & Funch-Jensen, 2004). Virtual Endoscopic Surgery Training (VEST), Minimally Invasive Surgical Trainer - Virtual Reality (MIST-VR), and MicroTEC.
Rehabilitation	Rehabilitation for patients with cognitive and motor disabilities (Deutsch, Latonio, Burdea & Boian, 2001).
Social situations	For the treatment of phobias (Pertaub, Slater & Barker, 2002) and for training social skills (Kerr, Neale & Cobb, 2002).
Underwater vehicle training	To Train operators to use remotely operated vehicles (Roberts, Pioch, & Ferguson, 2000).
Sports	Virtual hang gliding (Soares, Nomura, Cabral, Dulley, Guimaraes, Lopes & Zuffo, 2004). Location-Based Entertainment Pods and cockpits i.e., flight and driving simulators, and tennis (Molet, Aubel, Capin, Carion, Lee, Magnenat-Thalmann, Noser, Pandzic, Sannier & Thalmann, 1999).

Table 1.2: VE Applications (design).

Application	**Purpose**
Design	
Architectural design and city planning	To allow clients, the design team, and the public to view visualisations, and to conduct real-time walkthroughs, of proposed developments for evaluation before construction (Herder, Wörzberger, Twelker & Albertz, 2002).
Heavy engineering	To help the design and planning of manufacturing plants, for the petrochemical industry, marine, submarine, oil rigs, power stations etc. At present used, by Pepsi, BMW, ICI, Shell, Chevron, Fluor Corporation etc.
Aerospace industries	To help the design of aeroplane interiors (aesthetics and ergonomics) virtalis Ltd. To evaluate maintenance task procedures (Boeing Corp. Joint Strike Fighter maintenance feasibility study).
Automative	For the design of cars (manufacturing, aesthetics, and ergonomics). Allows designers to investigate and fully evaluate various designs before manufacturing, used by all the major car manufacturers.
Control rooms	For the assessment of ergonomic design, and working facilities for power stations and manufacturing plants.
Product design	For virtual prototyping of objects to gain an understanding about aesthetic and ergonomics.
Stress analysis	Finite Element Analysis (FEA) for objects and engineering structural analysis (Connell & Tullberg, 2000).
Computational fluid dynamics	Visulisation of fluid flows to help the understanding of design decisions (Chen, da V. Lobo, Hughes & Moshell, 1997).
Ergonomics	Human body stress, reach analysis etc. Space Station Freedom (NASA).

Table 1.3: VE Applications (data visualisation).

Application	**Purpose**
Data visualisation)	
TV weather forecast	For the communication of local and global weather patterns, also for solar wind and climate research, The Met Office, UK. National Center for Atmospheric Research, US.
Sports	Hawkeye (cricket) for the visualisation for cricket ball tradjectories for public broadcast.
Election and sports results	For public broadcast stations and general communication, for example VIZRT.
Medical data	Using information from CAT, MRI and PET scanners, 2D and 3D Visualizations of body organs can be created for diagnosis and treatment, for example IDL from Research Systems, Inc., and IMOD from the University of Colorado.
Scene of crime	For the analysis of a crime scene for investigators, Jury, and for training and performance monitoring, Advanced Interfaces Group, the University of Manchester.
Nano-engineering, scanning, and tunnel microscopy.	To help in the creation of new materials from the manipulation of individual atoms.
Geological and earth science data	To visualise and help the exploration of geological data captureed by satalite and direct sampling.
Simulations	To visualise the results of computational simulations, for example the Virtual Windtunnel (VWT) used by NASA to visualise and investigate the flow of air around the Shuttle.

Table 1.4: VE Applications (entertainment).

Application	**Purpose**
Entertainment	
Computer games	The popularity of PC and console computer games is having a growing impact on many peoples leisure time. The annual turnover of the computer games industry has seen an increase of 10 to 30 per cent per year every year since mid 1990. In 2002 £10.9 to £12.8 billion was spent worldwide on console games and equipment, £1.6 billion in the UK in 2001 (The Scotsman newspaper. Monday 30th September, 2002).
Location based entertainment/simulators	Theme park simulator rides create fantasy based theme park experiences. Sometimes inspired by scenes from Hollywood films e.g. “Back to the Future” by Universal, or from activities such as hand gliding, e.g. “Soarin’ Over California” by California Adventure, where the audience watches the action via an IMAX screen while sitting in ski-lift-type chairs that hang below a giant wing. Museums and art galleries have also employed VEs to educate and entertain, e.g. the immersive interactive works of Char Davies, Osmose (1995) and Ephmre (1998).
Cultural heritage and museum exhibits	VEs have been employed in the visualisation of historical landscapes and artefacts. They are used to reconstruct historical places and events, and for visualisation for archaeological exploration (Time Team, Channel 4).
TV reconstruction	The virtual reconstruction of situations, events, objects, and places for television bradcast.
Virtual tourism/marketing	VEs have been used to allow the virtual tourist to explore the Mars Pathfinder landing site (NASA), and virtual Stonehenge, virtual Notre Dame cathedral etc.

points and directions. For VE applications used for visualisation users also need to view the data from a variety of directions, while interacting to explore the effect of certain parameters. For entertainment applications, such as first person perspective computer games, navigating realistic or fantasy environments is often a fundamental part of the gaming experience. In fact, navigation is an integral part of a large number of VE applications. Navigation should be easy to perform so that users can move around and see any view with ease, while maintaining knowledge of their position and orientation. However, studies show that the opposite is often the case, with users often finding navigation extremely difficult even in VEs that have a fairly simple structure (Ruddle & Jones, 2001).

In 1965 Ivan Sutherland described a computer generated scene that would allow users to interact with virtual worlds (Sutherland, 1965). Today, the general public is more familiar with the term virtual reality, coined by Jaron Lanier. Those who work in the field, however, tend to use the term “virtual environment” to avoid any misconceptions of it necessarily being an attempt to recreate reality. A simple definition of a VE then is a computer mediated interactive environment. Some researchers take a broad view of VEs to include 2-dimensional (2D) windowed applications but here it refers to 3-dimensional (3D) representations of space that users can navigate around and view from different perspectives. However, if a virtual *reality* could be created, then people would perform in exactly the same way as they do in the real world because the two environments would be, by definition, indistinguishable from each other. It follows that people could navigate in a virtual reality as effectively as they navigate in the real world, something that is patently not the case in the VEs used today.

1.2 Virtual Environment Technology

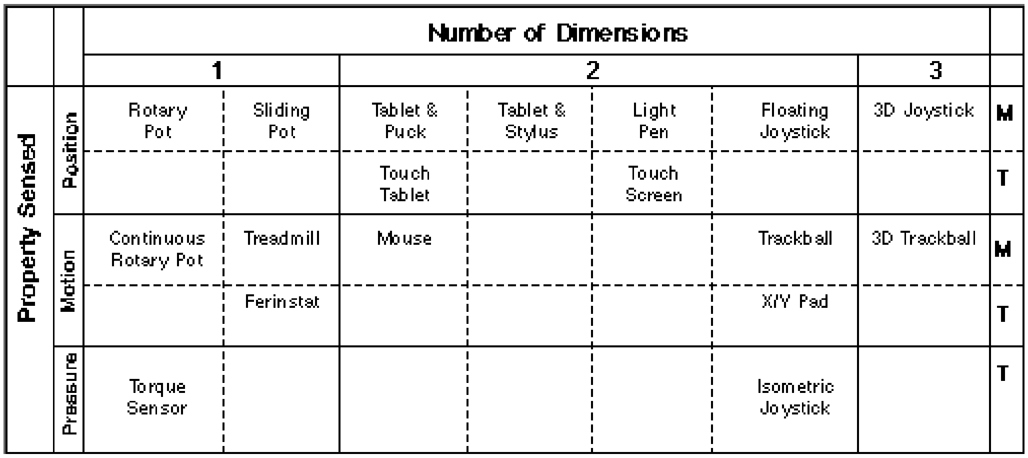
There are three general interactive elements to a VE application:

1. The input device(s) used for interaction within the VE.
2. The interaction technique (maps the movement of the device to effect change in the VE via the interface software).
3. The display device that communicate the result of the interaction.

1.2.1 Input Devices for Navigation

Input devices have a fundamental effect on users ability to navigate a VE because they influence movement by their physical design. Some input devices allow users to navigate a VE without the need to physically move their position in real space, others require some physical movement, and others require full body movement.

Rather than simply describe different input devices, it is useful to compare their properties using a taxonomy. Three of the best known taxonomies for input devices were presented by Buxton (1983), see Figure 1.1, Card, Mackinlay and Robertson (1991), see Figure 1.2, and Jacob and Sibert (1992).



		Number of Dimensions							
		1		2				3	
Property Sensed	**Position**	Rotary Pot	Sliding Pot	Tablet & Puck	Tablet & Stylus	Light Pen	Floating Joystick	3D Joystick	**M**
				Touch Tablet		Touch Screen			**T**
	Motion	Continuous Rotary Pot	Treadmill	Mouse			Trackball	3D Trackball	**M**
			Ferinstat				X/Y Pad		**T**
	Pressure	Torque Sensor					Isometric Joystick		**T**

Figure 1.1: The input taxonomy by Buxton (1983).

Buxton (1983) created a two dimensional graphical taxonomy showing: the number of spatial dimensions a device can sense (not to be confused with the number of degrees of freedom (DOF) that the device offers), and what action was being sensed, i.e. position, motion, or pressure. However, the Buxton (1983) taxonomy only included continuous devices, so Card, Mackinlay and Robertson (1991) developed their own taxonomy to include both discrete and continuous input, where discrete devices only measure one value and continuous devices are usually capable of measuring an ‘infinite’ (inf) number of values. In addition, Card et al. (1991) developed the concept of ‘composition operators’ to further describe the way various elements (i.e. buttons and sliders) are combined together to create the input device. For example, they describe the keyboard as a layout-composed device with approximately 100 keys on the same device. However, these taxonomies tend to treat devices that output the same information as equivalent and ignore the subjective qualities of the device. Jacob and Sibert (1992) added an important factor to the two taxonomies above by considering two components of interaction: the physical

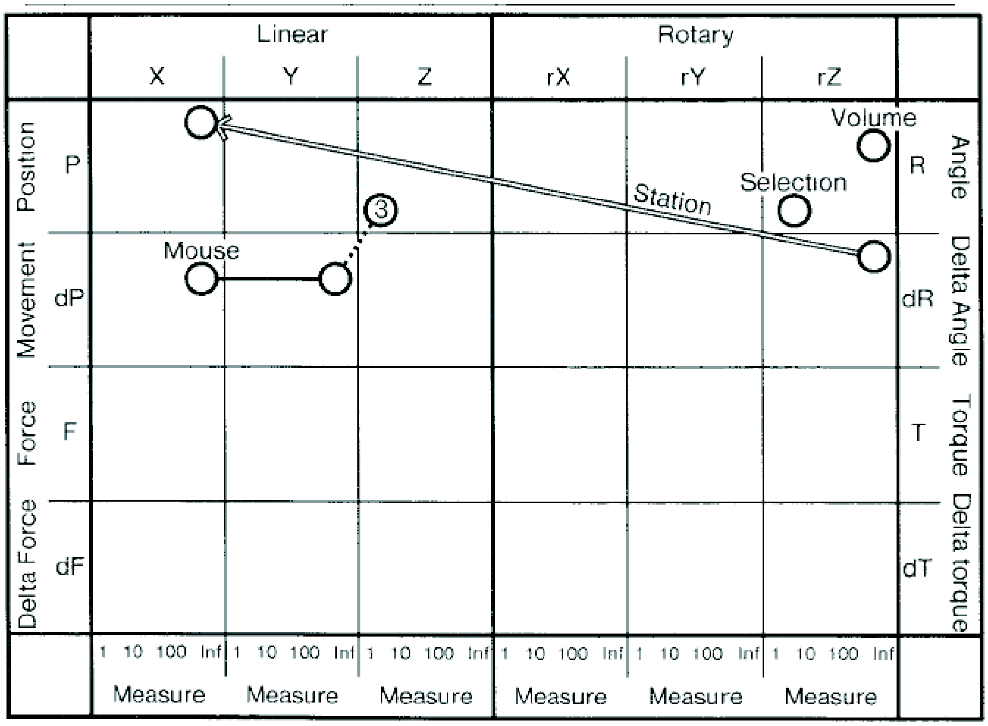


Figure 1.2: The input taxonomy by Card, Mackinlay and Robertson (1991).

properties of the input device, and the perceptual structure of the task space. In other words, how the task is perceived plays a crucial role in how the input device is used to achieve that task. Jacob and Sibert (1992) asked participants to move a 2D square, with either a mouse or a position tracker, to match the position of a target square and to either change it's size (seen as perceptually connected to moving its position: when objects are further away they appear smaller), or its shade (seen as perceptually unconnected to its position). Jacob and Sibert (1992) suggested that participants performed the movement and size change task quickest with the tracker because they perceived the task as three *integrally* related actions (move the square to match the position and size of the target square). However, participants performed the movement and shade change task quickest with the mouse because they perceived the task as two *separably* related actions (move the square to match the target position and then use the mouse button to change its shade).

The taxonomies above consider input devices generally. During the present PhD, a new taxonomy has been developed that focuses specifically on input for movement within a VE (see Figure 1.3). One axis presents the degrees of freedom, the second shows whether the input is discrete or 'infinitely' variable, and the third axis describes the type of movement input i.e. from the movement of the fingers or whole body. The most appropriate input devices for travel have been placed within this new VE travel taxonomy and are described in the rest of this section.

Keyboard: The keyboard is one of the two most common movement input devices for VEs, largely because it is one of the standard input devices supplied with personal computers. A single key can only be in one of two states, depressed or released, and so can only be used to control movement in one direction. Control of movement in the positive and negative direction of a single DOF requires the use of two keys. Taken as a whole, a keyboard is a layout-composed device (Card et al., 1991) that comprises approximately 100 keys.

Mouse: The mouse is the other very common interface device, and is another example of a layout-composed device. Two or three buttons (each has one DOF) are combined with a sensor that measures movement in the x and y directions across the horizontal plane (two DOFs).

Joystick: The joystick was first used to control aeroplanes, but is now used in a wide variety of situations for controlling movement in VEs. Most computer joysticks measure input in terms of two translationary DOFs and, in addition, buttons or other switches are usually attached to the joystick for discrete input making it another layout-composed device.

Spacemouse: (see Figure 1.4) The spacemouse is one of the few desktop input devices specifically designed for 3D VE interaction. The user applies pressure to, or twists what is essentially a short joystick, to move or turn objects or their perspective in the VE. Most joysticks pivot at their base when pressure is applied, but the handle (ball or cylinder) of a spacemouse typically moves very little and responds to very light forces. The handle of the spacemouse can detect rotational movement about three axes, and the nine buttons arranged around the handle can be used for other tasks.

3D mouse: (see Figure 1.5) A 3D mouse is like a conventional desktop mouse in that it combines the measurement of movement with four buttons and a joystick into yet another form of layout-composed device. However, unlike its conventional counterpart, a 3D mouse measures position and orientation in 3D space (6 DOF) using a tracking device such as the Ascension Flock of Birds (see below).

Trackers: Motion trackers are small devices that track their position and/or orientation in space. They can be attached to hand held devices, such as a 3D mouse, or to displays such as the HMD to track rotational and translational movements over an extended area (e.g. 10 x 10 metres). Motion trackers use a number of technologies including:

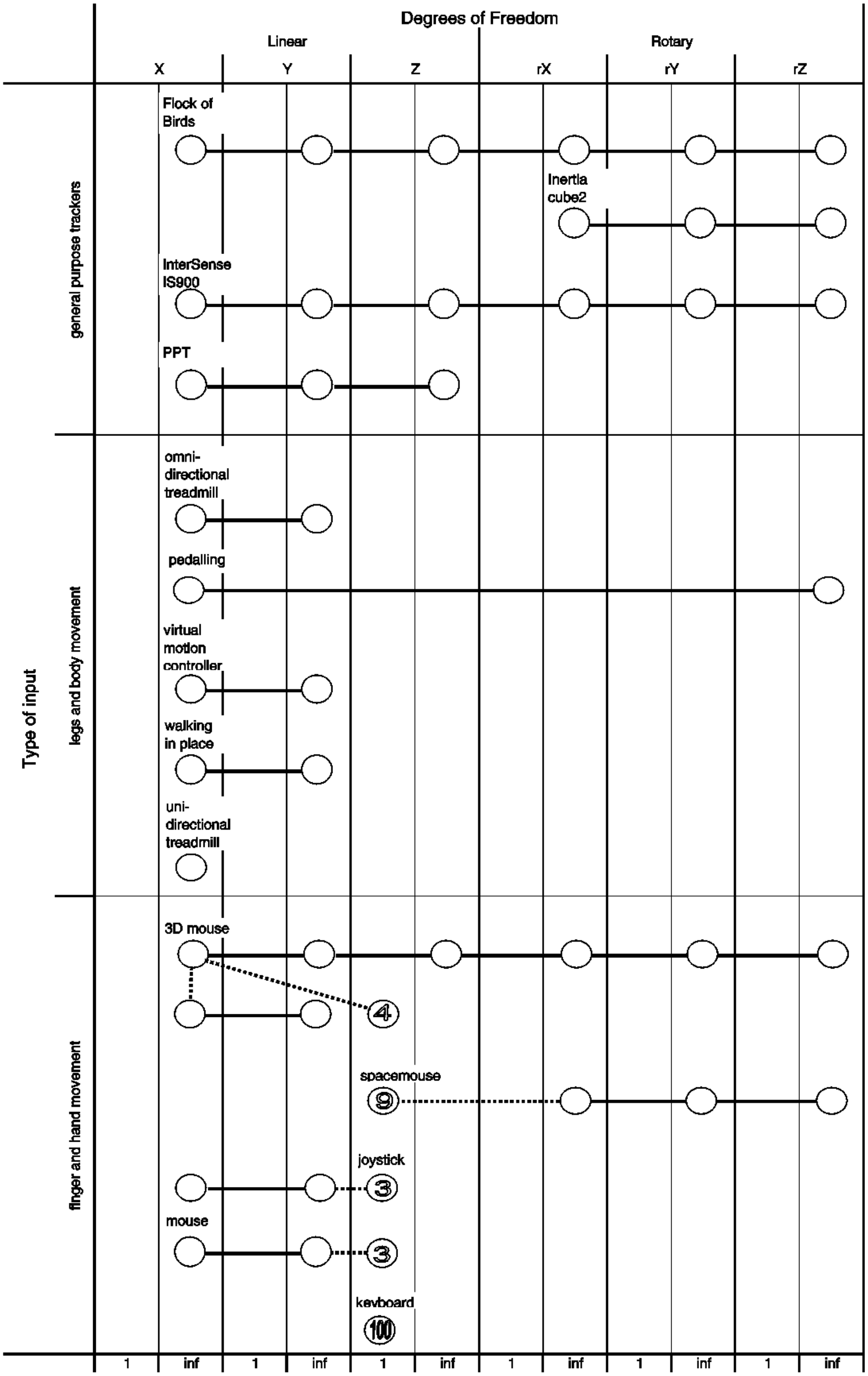


Figure 1.3: A taxonomy of input devices for travel in VEs.

Figure 1.4: The Spacemouse.

Figure 1.5: The 3D mouse.

Mechanical. Mechanical trackers are limited by the reach of their arm, generally, the larger the reach the greater the cost. The main advantages of mechanical tracking are that there is low latency and it is highly accurate, it therefore lends itself to object manipulation and other dexterous tasks.

Magnetic, e.g. the Ascension Flock of Birds, use low frequency magnetic field emitters to pick up the position and orientation of a small sensor (receiver) that is attached to the tracked device. They can operate within a volume of approximately 1.5 to 10 metres but the tracked space is subject to distortion from metal objects located near the sensor and they tend to be associated with noticable latency.

Inertial, e.g. the InterSense IS900 (position and orientation), and the InterSense Inertial Cube2 (orientation only). Inertial trackers use small angular-rate gyroscopes and detect rotations. The sensor is attached to the tracked device (3D mouse, HMD etc) and can suffer from error accumulation when used for long periods. However, trackers like the IS900 use sonic beacons to correct for these errors. Also, unless a wireless transmission system is used the tracked volume is limited to the length of its connecting cable.

Optical, e.g. the World Viz Precision Position Tracker (PPT). Optical trackers use two or more cameras placed around the environment that detect, via computer vision software, the position of either a light emitting diode (LED), or reflected infrared light from fiducials (small reflectors). Optical trackers can track a larger volume than can mechanical or magnetic trackers, but require line of sight to be maintained between the cameras and LED/fiducial so triangulation can be performed to calculate position.

Virtual Motion Controller (VMC) The VMC (Peterson, Wells, Furness & Hunt, 1998) is a disc that pivots in place as the user stands and leans on its edges. The disc tilts slightly with the weight of the user and this movement is used to signal in which direction and at what speed to move the user in the VE. This device is designed to employ the legs and requires some physical movement to operate and therefore provides some proprioceptive feedback to the user while leaving the hands free for other tasks.

Pedalling Various movement input devices have been used as a method to travel around the VE that require participants to pedal, such as a uni-cycling (Darken, Cockayne & Carmein, 1997), or bicycle (Carraro, Cortes, Edmark & Ensor, 1998; Harris et al., 2002), or tricycle (Allison et al., 2002). These devices, like the VMC above, require physical movement to operate which provides proprioceptive feedback from pedalling and steering, but lack vestibular feedback from actual physical movement.

Uni-directional treadmill: Treadmills are locomotion devices that allow users to walk without changing location (Mohler et al., 2004). However, because the surface that users walk on allows input on only one axis movement in the VE is somewhat restricted. With the uni-directional treadmill the VE is usually projected onto a large screen in front of, or sometimes surrounding, the user, or displayed via a HMD.

Omni-directional treadmill (ODT): (see Figure 1.6) This device is a platform constructed from small rollers, the user walks in place on the horizontal plane in any direction which moves the user's perspective through the VE (Darken et al., 1997). Of all the input devices discussed so far the ODT is the one that most closely resembles actually walking. Like it's uni-directional counterpart the VE is usually projected onto surrounding screens or displayed via a HMD.

In addition, an alternative to the ODT is "walking in place" (Slater, Usoh & Steed, 1995; Templeman, Denbrook & Sibert, 1999). In this a users physical movement is tracked using devices like a Flock of Birds, and the motion interpreted by software as a form of walking. The result is movement with the same number of DOFs as the ODT.

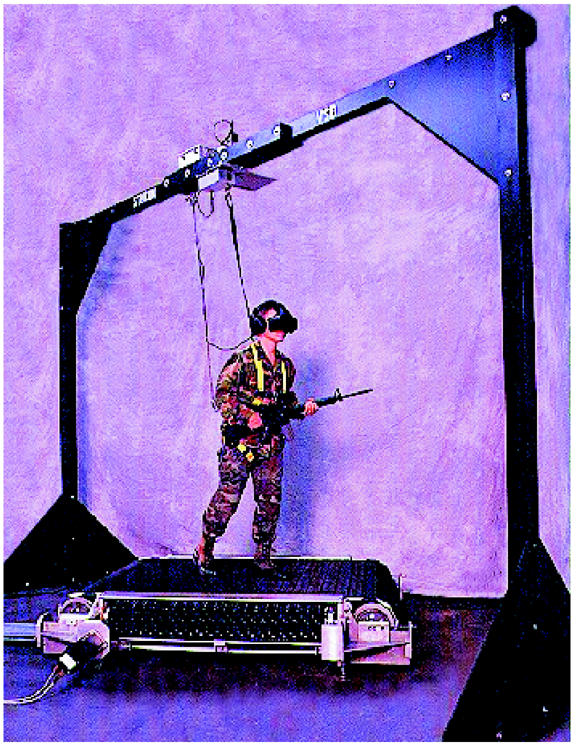


Figure 1.6: The omni-directional treadmill (left, courtesy of David Carmein of Virtual Space Devices).

1.2.2 Interaction Techniques for Travel

In a VE application travel is usually instigated by the manipulation of an input device with the resulting movement controlled by the implementation of the software. How that software controls those input signals is described as the interaction technique.

Bowman, Kruijff, LaViola and Poupyrev (2004) and Bowman, Koller, and Hodges (1997) describe several different methods to classify interaction techniques for travel. One such method, task decomposition, subdivides the travel task into three subtasks:

1. Direction or target selection, that is, how the direction of travel, or the place where one wants to travel to, is selected. For example, the user can travel in the direction of their present view (gaze directed steering), discussed below, or a target location can be selected and the software can move the user's viewpoint to that location.

2. Velocity/acceleration selection, which refers to how users control their speed, e.g. whether they can choose constant or variable velocity.

3. Conditions of input, which refers to how the input is controlled, e.g. continuous velocity while a trigger is depressed and zero velocity when released, or perhaps an on/off switch controls the speed requiring no further input until the user wants to stop.

Interaction techniques can also be classified by movement metaphors that are used to aid users familiarity while navigating. Poupyrev, Weghorst, Billinghurst and Ichikawa (1998) classified interaction techniques as either exocentric (i.e. the user interacts as though looking into the VE from outside) or egocentric (the user interacts as though inside the VE).

Bowman et al. (2004) organised travel techniques into six metaphors: steering, route-planning, target-based, manual manipulation, scaling, and physical locomotion. The steering metaphor is used to liken the movement of travel to that of a car or similar vehicle. Often, users input their speed with one control (e.g. mouse or keyboard button) while they steer with another (e.g. a steering wheel or mouse). Route-planning refers to the ability to plan a route through the VE before travel. This technique can be used with an exocentric view to define and edit a path in relation to the VE as a whole. With target-based travel techniques users point to, or otherwise indicate, where in the VE they want to travel. The VE software then moves, often along a continuous path, to that target. Manual manipulation techniques such as “grabbing the air” can be used to move the user’s viewpoint around the VE. This technique may be used with one or two position-tracked hands as though pulling on a rope. Travelling by scaling refers to the manipulation of the relationship between the users physical movement and the resultant movement in the VE. By scaling up the user’s physical movement a user can travel around a larger VE and stay within the, often limited, tracked space. This technique can be used with other metaphors such as “grabbing the air”, one movement of the hand could, for example, move the user one or 20 metre etc. Lastly, the most natural way to travel in a VE is to physically walk around it. Physical locomotion can be tracked to manipulate the user’s viewpoint in the VE. If the VE is larger than the tracked space a user’s movements can be scaled to allow travel to all parts of the VE.

Ware and Osborne (1990) evaluated the appropriateness, through interviews with participants, of three interaction metaphors for navigation and virtual camera control. The metaphors used were *eyeball in hand*, *Scene in hand*, and *flying vehicle control*. They found that no one metaphor was judged best for all tasks; each technique had its advantages and disadvantages.

While travelling through a desktop VE, across a horizontal plane, there are three main directional elements to movement: the direction of a person’s view, the orientation of their body and their direction of travel. Altering the relationship between these three elements creates the three primary walking metaphors used for travel within virtual environments (Ruddle & Jones, 2001). The first of these is view-direction (gaze-directed) travel (Bowman, Johnson & Hodges, 2001; Bowman, Koller & Hodges, 1997) where the heading of the body, the direction of travel, and the direction of view are all locked together; the user can only travel in the direction they are looking. The second is body-direction travel where the heading of the body and the direction of travel are locked together, but one can manipulate the direction of view independently of the other two (i.e. the view and travel directions are decoupled). Lastly there is independent movement where one can travel independently of both the viewing direction and body direction. Independent movement is the method that most closely resembles natural human movement. As one progresses from view-direction, to body-direction, and then independent movement, there are an increasing number of degrees of freedom (DOFs) available to the user.

Different interaction techniques can be implemented with the same input device. For example, the keyboard is often used to control the speed of travel in a desktop VE, a depressed key can either signal the start of constant movement or the start of acceleration up to a maximum. Also, the movement of the mouse is often used to vary the direction and elevation of the user's view. Moving the mouse left and right usually turns the view direction left and right (heading/yaw) i.e. there is a stimulus-response compatibility. However, increasing the elevation of view (pitch) can either be controlled by moving the mouse away from the user, used for most ground based VEs, in which case the mouse is being used with a pendulum metaphor that is turned 90 degrees up towards the screen, or by moving the mouse towards the user, as in most desktop flight simulators, in which case the mouse can be thought of as the top of a joystick.

Many different types of interaction technique have been developed, along with a large number of combinations of technique and device. The effects that the most important of these have on navigation are discussed in Chapter 2.

1.2.3 Display Devices

The user interacts with a VE via various physical input devices and the interaction technique interprets those signals to create movement within the VE. The next section discusses the third part of that process, displaying the result to the user via various visual displays.

Taxonomies of display devices are far less common than they are for input devices, but one relevant to navigation has been developed for this research (see Figure 1.7). The dimensions included are whether a display is space constant or consistent (Zheng, McConkie & Schaeffer, 2003), the FOV provided and number of viewers that maybe accommodated. In terms of cost, display devices can range from less than £100 (desktop monitor) to hundreds of thousands of pounds (CAVEs).

Desktop monitor: The desktop monitor is the commonest display type for nearly all computer generated environments, primarily because of its availability and low cost. Zheng et al. (2003) describe the desktop monitor as a space-constant display device. That is, while travelling through a VE the user changes the position and orientation of the VE camera while the monitor and user remain stationary. Therefore, there will be no proprioceptive or vestibular feedback available to the user while travelling through the VE, unless a pedalling or treadmill input device is used.

Although monitors display monoscopically, they are capable of displaying stereoscopically with additional hardware. Stereoscopic viewing can either be achieved actively or passively. Active stereo glasses use fast shuttering LCD filters to allow only

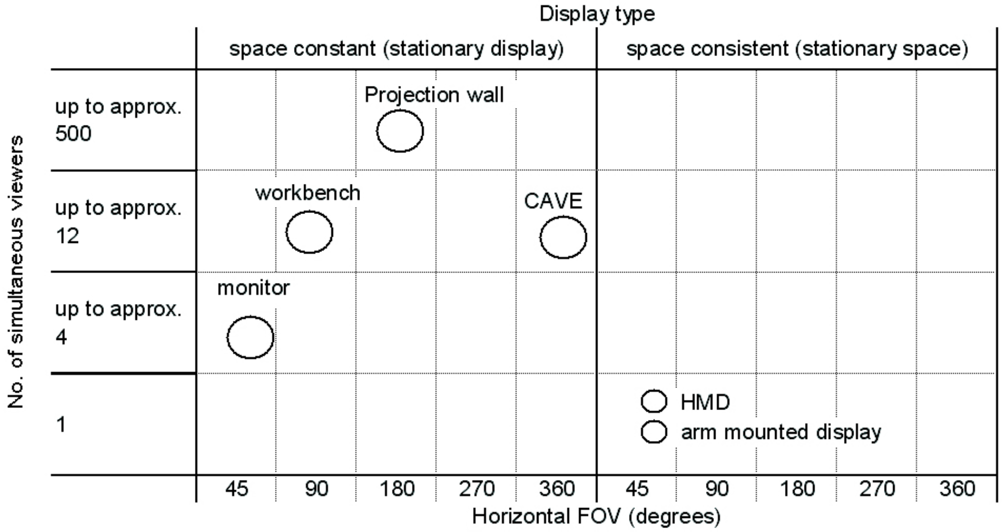


Figure 1.7: Display taxonomy.

the left eye to see images intended for the left eye, and the right eye to see images intended only for the right eye. The two sets of images are therefore displayed alternately, typically, 60 frames per second, 30 going to each eye. A passive stereoscopic display is achieved by polarization multiplexing, viewers wear glasses that contain two polarizing filters placed at right angles to each other, the display contains two oppositely polarized images so the lenses only allow through one set of images for the left and the other set for the right eye. Although stereoscopically displayed images are seen as more realistic and engaging there is little evidence so far to show that it offers any benefit for navigation tasks over VEs displayed monoscopically.

It is possible to combine tracking technologies (see above) with desktop monitors to provide what has been described as fish tank virtual reality (VR) (Ware, Arthur & Booth, 1993). The tracker records the position and orientation of the user's head and that information is then used to update the screen images, viewed actively or passively, to create motion parallax depth cues.

Workbenches: Display benches, like the TAN Holobench, usually provide one or two back projected planes of translucent plastic/glass and offer a viewing area considerably larger than a monitor. The display is often projected stereoscopically which makes them idea for displaying virtual objects to single or small group of observers. Head tracking can be used to create motion parallax, but like fish tank VR, only the tracked individual will be able to view the correct depth cues.

Projection wall: (see Figure 1.8) the VE can be projected, using one or more projectors, onto a smooth white surface (curved or flat) to create a large display.

The images may be projected from the viewer's side, similar to a cinema screen or IMAX theatre, or onto the reverse side of a translucent surface, as in the Holobench (see above). Wall projections usually cover a large surface area and can provide large groups of observers with a wide FOV, though rarely one of more than 180 degrees.

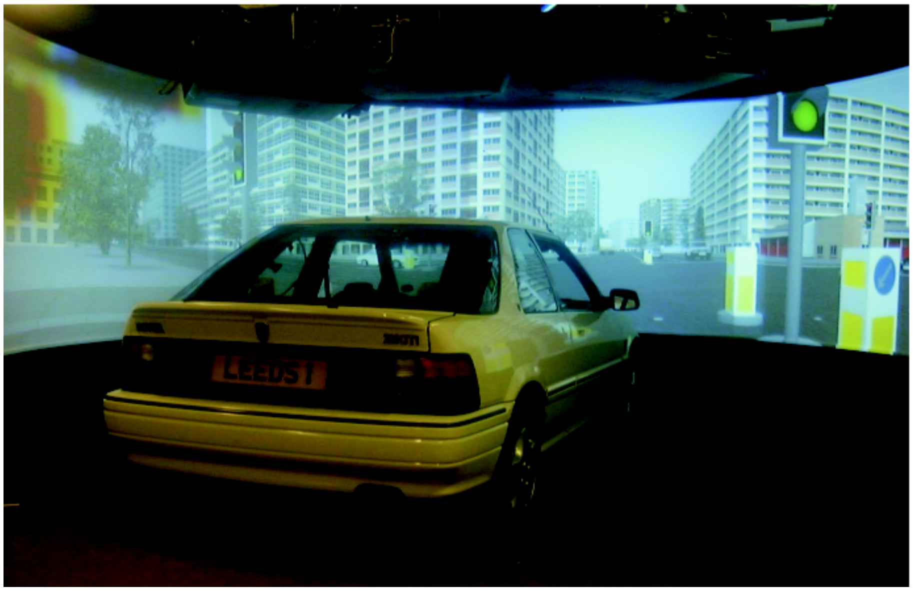


Figure 1.8: The projection wall used for the University of Leeds driving simulator.

CAVE: The CAVE (CAVE Automatic Virtual Environment; Cruz-Neira, Sandin & DeFanti, 1993) utilises four to six translucent screens arranged in a cube configuration that surrounds the user(s), the VE is then projected onto the screens from the outside and fills the users' FOV. Users can physically move around in the CAVE that with head tracking can create motion parallax cues, but the surrounding screens limit physical movement. Both projection screens and CAVEs represent a significant investment thereby limiting their accessibility, and both types also require considerable space.

HMD: (see Figure 1.9) The head mounted display provides two miniature displays in front of the users eyes. Zheng et al. (2003) describe the HMD as a space-consistent display because when head rotation are tracked the VE appears to remain stationary while the user turns to look around. HMDs are associated with increased incidences of simulator sickness when compared to the fixed screens of monitors and projection screens therefore limiting accessibility. In addition, HMDs are not particularly comfortable to wear, not only are they relatively heavy (the Virtual Research V8 HMD weighs approximately 1 kg), but the centre of gravity is often at the front and tips the device forwards on the head. A counter-balance can reduce the problem but this also increases the weight of the device. HMDs are, however, one of the most immersive display technologies available and as such can produce an engaging sense of presence. HMDs can, with the use of wide area tracking, be used to view a VE while simultaneously walking around real space, thus creating the illusion of walking around a VE (see tracked walking above).

Figure 1.9: The Virtual Research HMD v8.

Arm Mounted Display. The arm mounted display (such as the BOOM, Binocular Omni-Orientation Monitor; Bolas, 1994) consists of an armature with a position and orientation tracked binocular display device (similar to an HMD) at one end and its counterbalance at the other. Although the stereoscopic display device can only be manoeuvred within an area accessible by the arm, the arm and counterbalance reduces the weight carried by the user and this allows heavier and higher quality components to be used than would be practical with a HMD.

1.2.4 Rendering Technology

Approximately 10 years ago most VR systems ran on Silicon Graphics Inc (SGI) Reality Engine graphics supercomputers (or other similar image generators) and typically cost tens or hundreds of thousands of pounds. Despite their high cost they were limited in their ability to render advanced graphics such as texture or bump maps, surface reflections, or even some shade and shadow effects. A PC a decade ago (with 4-8MB RAM and the newly introduced 3D acceleration cards) could offer considerably less graphic rendering capabilities.

A standard PC today, however, has the ability to produce far superior graphics than the SGI machines could a decade ago and for a fraction of the cost. New rendering techniques, such as vertex/pixel shading and texture mapping, greatly improve the visual realism of VE scenes by adding shade and shadows and surface textures to the geometry implying scale and surface material. Advances in the speed of PCs allow many advanced graphics techniques, such as surface reflections, to be calculated in real-time.

1.3 Summary

Participants acquire spatial knowledge from VEs substantially slower, and less accurately, than that acquired from the real world. Navigation is a fundamental activity in most 3D VEs. Without the ability to navigation accurately participants will not be able to realise the potential of VEs for the acquisition of spatial knowledge. The overall goal of this thesis is to study why people have difficulty navigating in VEs.

The taxonomies discussed, and the new taxonomy created, help to compare the various VE input and output devices in a common design space. This comparison potentially allows a better understanding of the three VE system variables (FOV, visual fidelity, and movement interface) and their influence on participants' ability to navigate. Chapter 2 discusses the existing research literature on the three VE system variables, and proposes three levels of metric to help VE task analysis. Chapter 3 provides an overview of the four experiments conducted. The chapter considers issues such as materials and procedures common to all of the experiments, and eliminates the need for repeated experimental data. Chapter 4 describes a real world experiment conducted to provide comparable data for the three subsequent VE studies. Chapters 5, 6, and 7 describe the experiments that investigated the effects of FOV, visual fidelity, and movement interface on participants' ability to navigate while conducting a VE search task. Chapter 8 provides sickness data on the VE studies. The thesis concludes with a general discussion on the issues raised.

Chapter 2

Navigation Research

This chapter briefly discusses the different categories of space that are used in navigation. It then reviews existing research conducted on each of the three aspects of a VE application outlined in the introduction (see Chapter 1), namely:

1. Field of view (FOV, graphical and physical), in other words, how much of the visual scene can be seen at any one time.
2. Visual fidelity. How closely the VE scene resembles its real world counterpart.
3. Movement interface. The use of various input devices and techniques for travel around the VE.

The research literature discussed in this chapter provides an overview of present knowledge concerning participants' performance and behaviour for navigation tasks in real and virtual spaces. It is also intended to show how the experiments described in this thesis augment existing knowledge, and address the question of why participants do not generally navigate VEs as well as they do real world environments. The chapter concludes by considering different metrics that are used in present VE navigational research, and reviews a more structured approach that considers three related categories of metric to evaluate VE applications.

Weatherford (1985) identified three categories of space that are used in navigation. Model-scale spaces are those that can only be observed from the outside, for example, a model displayed on a tabletop. Small-scale spaces are those where it is possible to see the entire space simply by standing in one place and looking around, examples include single interior rooms of buildings or external courtyards. Large-scale spaces are those in which it is not possible to resolve all the details necessary for efficient navigation from a single position. Large-scale spaces are comprised of several smaller spaces and to experience the extent of the large-scale space one has to navigate and change viewpoints. Examples include buildings containing several spaces occluded from each other by walls and doors, urban areas of a city, and extended land and seascapes. The studies described in this thesis all used small-scale real-world and virtual spaces.

2.1 Field of View (FOV)

The effects of FOV have been studied for both real world and VE spatial tasks. In one well known real-world study (Alfano & Michel, 1990), participants were asked to reconstruct the layout of a real space viewed with a FOV that ranged from normal to nine degrees. The restricted FOV not only distorted participants' perception of the size of the space, but also reduced their ability to accurately reconstruct the spatial layout using colour copied photographs. However it was only with a very narrow FOV (22 degrees or less) that participants' performance was significantly worse than with a normal FOV. A more recent study investigated egocentric distance perception in the real world using a visually directed walking task (Creem-Regehr, Willemsen, Gooch & Thompson, 2003). Participants were first shown a target and then had to walk to its position while blindfolded. Participants who used a restricted FOV (42 x 32 degrees) were just as accurate as those who viewed the target with a normal FOV. However, simultaneously eliminating head rotations by using a neck brace and restricting the FOV produced systematic underestimation of the distance to the target, but the reason for this remains an open question.

Studies that require participants to estimate distances and perform other spatial tasks in VEs often suggest that the restricted FOV contributes toward poor spatial performance. For example, when participants judged egocentric distances while wearing a HMD, they underestimated the distances while using a wide FOV (140 x 90 degrees) but overestimated the distances while using a narrower FOV (60 x 38.5 degrees; (Kline & Witmer, 1996). More recently, Czerwinski, Tan, and Robertson (2002) showed that with a novel navigation technique (the faster participants moved forward the higher and steeper they viewed the VE) males performed better than females with a narrow FOV (desktop display), but when the FOV was widened females performance increased to equal that of males. However there are also VE studies that have shown no influence of FOV for spatial tasks. In one, participants performed a series of triangle completion tasks using a large projection screen, and accuracy was influenced by path layout but not by FOV (Péruch, May

& Wartenberg, 1997).

Another study investigated participants' ability to estimate ego-rotations from optic flow (Schulte-Pelkum, Riecke, von der Heyde & Bülthoff, 2002). Participants were instructed to turn five angles that ranged from 45 to 225 degrees while they viewed a VE that consisted of a star field. Each star had a limited life of 650ms which limited participants' ability to use landmarks. There was no differences between participants estimations of turned angles for the two FOVs presented on a projection screen (86 x 64 degrees, and 40 x 30 degrees), but with both of these participants were significantly more accurate than when the judgments were performed using a 40 x 30 degree HMD.

Why was the use of a screen, with a FOV 40 x 30 degrees, more accurate than the HMD with the same FOV? An explanation might possibly lie in the perceived size of the space with which participants performed the task and its resultant effect on participants' strategies. Tan, Gergle, Scupelli, and Pausch (2004) looked at different sized screens that subtended the same visual angle, i.e. the larger screen (1.95m x 1.5m) was further away than the monitor (0.36m x 0.27m). They found that, for a triangular path completion task, participants who used the physically smaller monitor underestimated distances significantly more than those participants who used the large projection screen. They suggest that the larger screen encouraged participants to adopt a more efficient egocentric strategy that then produced more accurate results. The previous studies show that both PFOV and display mode seem to play their part in participants ability to accurately estimate distance and angles turned; two subtasks that are performed while navigating a VE.

Desktop VEs are displayed via monitors. For a normal viewing distance (600 - 800mm), this gives a physical field of view (PFOV) of approximately 48 degrees (horizontal) by 36 degrees (vertical), which is substantially less than the normal human FOV of 200 x 160 (May & Badcock, 2002).

The PFOV is of course fixed, unless a physically larger monitor is used, but the geometrical field of view (GFOV) of the VE camera can be varied. The smaller the GFOV the less VE can be seen at any one time (similar to looking through a telephoto lens), while more of the VE can be seen with a larger GFOV. However, displaying large GFOVs (say 100 degrees plus) creates considerable visual distortion of the VE, the effect of which is similar to viewing the VE through a wide-angle (fish-eye) lens. Of course, mismatches between the PFOV of a display and the GFOV of a VE are not the sole preserve of desktop VEs, and have been deliberately applied for navigation tasks with other displays such as HMDs (Ruddle, Payne &

Jones, 1999).

Waller (1999) investigated participants ability to accurately estimate exocentric (interobject) distances in a VE via a desktop monitor (horizontal PFOV = 50 degrees) with different GFOVs and either with or without error-corrective feedback. With feedback participants tended to underestimate distances with a 50 degree GFOV, were more accurate with a 80 degree GFOV, and overestimated distances with a 100 degree GFOV. Without feedback, however, all GFOV conditions were overestimated. This suggests that the GFOV is a significant factor in influencing participants' exocentric distance perceptions but repeated corrective feedback on accuracy can inform, and therefore allow correction by, participants on the distortions produced by different GFOVs.

Even when the GFOV exactly matches the PFOV there can be considerable distortion in peoples perception of a VE. In an experiment by Psotka. Lewis, and King (1998) participants viewed a monitor that displayed a VE of the room in which they sat. As the virtual camera slowly rotated the observers were asked to draw on a plan of the VE the path of the animated camera. Even though the virtual camera rotated in place and did not change position, participants consistently represented the travelled path as an ellipse; participants thought that the virtual camera moved much closer to the objects in the VE than it actually did. However, the introduction of deliberate spatial distortion, in the form of the implementation of a corrective wide-angle view, can correct for this type of bias (Ellis, 1993).

Attempts have been made to represent the full 200 degrees of the human FOV onto display devices to increase users' awareness of the VE. Slater and Usoh (1993) used images from five virtual cameras and displayed them via an HMD in a picture frame configuration. The middle 'picture' displayed the viewer's normal (forward) view, with no significant distortions, while the outer 'frame' displayed compressed views from the other four virtual cameras. The outer frame of the display gave the viewer an idea of what was directly to the left, right, above, and below the view ahead. Unfortunately, this study did not attempt to analyse the possible benefits to participants engaged in a spatial task. However a study by Robertson, Czerwinski, and Dantzich (1997) employed the same idea, this time on a monitor and only on the right and left of a main window. They asked participants to visually search for a target letter while travelling through a 3D VE consisting of four corridors all 18 metres in length. They found that, when the undistorted main window remained a consistent size, participants found the target letter no quicker with the peripheral lenses than without.

A restricted FOV has been identified as a hindrance to collaborative tasks within a VE. In one experiment Hindmarsh, Fraser, Heath, Benford, and Greenhalgh (2000) asked either two or three participants to collaborate on a virtual furniture arranging task. They noted that the limited FOV (55x45 degrees) meant that participants could not simultaneously see the object they were manipulating and their co-participant(s). This resulted in increased compensatory verbal instructions to explicitly communicate what would normally be done so by participants' implicit actions. Interestingly, they also implemented peripheral lens in an attempt to compensate for the lack of FOV but provide no results on performance.

In summary, there is conflicting evidence as to whether a modest reduction in participants FOV (e.g. to 50 degrees) substantially effects performance in spatial tasks. Participants tend to overestimate VE distances with a narrow FOV and underestimate them with a wide FOV (Kline & Witmer, 1996). However, not all have found effects of a restricted FOV (Creem-Regehr et al. 2003; Péruch et al. 1997; Schulte-Pelkum et al. 2002; Tan et al. 2004; Psotka et al. 1998). Attempts to use display techniques on a desktop monitor to compensate for the lack of peripheral vision have, so far, been unsuccessful (Slater & Usoh, 1993; Robertson, et al. 1997).

2.2 Visual Fidelity

Waller, Hunt, and Knapp (1998) introduced the concept of environmental fidelity, that is, how closely a VE resembles its corresponding real world scene. Environmental fidelity has many different factors, including the structure of an environment, its visual characteristics (e.g., whether every real-world object is included in a virtual scene, and the detail with which each object is modelled), and other sensory information.

Visual characteristics are one of the primary sources of information that people use to determine their position and orientation within an environment. The role of visual information is particularly important in VEs, where there is often no non-visual sensory information. It follows, that changing the amount of visual information available will affect users behaviour and that will, in turn, affect their ability to perform spatial tasks in a VE.

Presence, the subjective sense of being in a VE, has been seen as an important aspect of VE research for many years. However its relevance and link with task performance is still in debate. In one study, Welch, Blackmon, Liu, Mellers and Stark (1996) investigated the role of interactivity (active vs. passive control), pictorial realism (high vs. low visual detail), and delay of visual feedback (220 milliseconds vs. 1.5 seconds), on participants sense of presence for a driving task. Results showed that, on the reported sense of presence, active control, high visual detail, and low latency produced a higher sense of presence than the other three, but visual realism was judged as less important when compare to the other two. Unfortunately, metrics of participants' performance were not reported. As Welch et al. (1996) acknowledge, there is an inherent difficulty in investigating the link between presence and performance because aspects that increase one are also likely to increase the other.

The majority of studies that have investigated VE navigation have deliberately used bare and simplistic virtual scenes, albeit often texture mapped. However, a notable exception was a study that compared the acquisition of route knowledge from a high fidelity VE building with other training media (Witmer, Bailey, Knerr & Parsons, 1996). This found that participants who were trained in a VE made more navigational errors than those trained in the real world, but fewer errors than participants who trained by being shown pictures.

Two attributes of the visual scene that have been shown to be of particular importance for spatial tasks are texture and landmarks. Textures have been shown to affect spatial tasks in three different ways. First, the textures enhance optic flow, which enhances users' perception of motion and has been shown to aid navigation (Kirschen, Kahana, Sekuler & Burack, 2000). However, if a user physically walks around a VE instead of using an abstract means of movement then optic flow can be reduced to a negligible amount without affecting performance on simple spatial tasks such as path integration, indicating a dominance of physical movement over visual information in certain forms of navigation (Kearns, Warren, Duchon & Tarr, 2002).

Second, repetition (tiling) of textures such as a brick pattern conveys metric information that significantly improves participants' accuracy when estimating distances (Sinai, Krebs, Darken, Rowland & McCarley, 1999). However other types of textures such as grass, carpet and abstract patterns had no effect on the accuracy of participants' distance estimations (Sinai, Krebs, Darken, Rowland & McCarley, 1999; Witmer & Kline, 1998). This suggests that certain types of textural information can be valuable for maintaining one's spatial orientation in VEs, particularly in the absence of useful landmarks and proprioceptive information.

Third, texture mapping has been used for many years as a cost effective method for improving the visual realism of scenes, for example, for building facades, trees, and signs. These types of textures often either are, or contain, distinctive "objects" that act as landmarks and could also be modelled as 3D geometry. The availability and type of landmarks are well known to influence participants search performance in VEs. For example, Steck and Mallot (2000) showed that participants stored both local and global landmarks in memory during a VE familiarization phase, and relied on one for navigation when the other was removed. The saliency of landmark cues has also been shown to influence navigational performance. In this respect, participants' route-finding accuracy has been shown to increase when everyday (i.e., familiar) objects were used as local landmarks, but not when coloured patterns were used instead (Ruddle, Payne & Jones, 1997). Tlauka and Wilson (1994) found that landmarks were useful for wayfinding when other strategies, specifically counting left and right turns, were suppressed by an artificial increase in workload (backward counting).

Although there are no common metrics that can be applied to measure the overall content of a scene for the availability of navigational cues, research has been conducted on the notion of scene complexity. The suggestion here is that the more complex the visual scene the more navigational cues will be available for the updating of self-position and orientation within a VE. In an experiment by Oliva, Mack, Shrestha, and Peeper (2004) participants were asked to hierarchically group 100 colour images of indoor scenes in order of visual complexity. One group was told that visual complexity related to how difficult it would be to describe the scene by a verbal description, while the second group was told that complexity related to how difficult it would be to make sense of the structure of the scene. The first group ordered the images largely based on the quantity and variety of objects and colours, while the second group ordered the images based on the organisation of the spatial layout of the image, i.e. clutter, symmetry, and open space. It appears that the first group made their judgments based on the scene that the image depicted whilst the second made their judgments based on the configuration of the image. Oliva et al. (2004) suggest that because the hierarchical grouping order differed significantly between the two groups of participants, complexity was best described by a multiple set of perceptual dimensions including the number of objects, colour, clutter, symmetry, open space, and organisation. All of which could potentially be useful for differentiating one location in a VE from another.

Finally, the research cited in this section suggests that the amount of visual information contained within a scene can affect participants' ability to perform spatial tasks in a VE. Considering that real world scenes usually contain a great deal of visual information, a VE that replicates a real world scene is likely to contain a similar amount of orientation cues given the same FOV. However, even with a VE visually similar its real world counterpart real world navigation performance is seldom obtained ((Witmer, Bailey, Knerr & Parsons, 1996; Sinai, Krebs, Darken, Rowland & McCarley, 1999; Witmer & Kline, 1998)).

2.3 Movement Interface

Our ability to move around a virtual space directly influences our perceptions of that space. The design of the movement interface, and a users' proficiency at using it will, therefore, directly influence their ability to perform spatial tasks. In fact, proficiency with the interface, measured by timing participants' performance on various simple navigational tasks in a VE maze, has been shown to be one of the most important factors affecting individuals' ability to perform spatial tasks in large-scale VEs (Waller, 2000).

As stated previously, there are three main directional elements to movement in a VE: the direction of a person's view, the orientation of their body, and their direction of travel (Ruddle & Jones, 2001). Altering the relationship between these three elements creates view-direction (gaze-directed) travel (Bowman, Johnson & Hodges, 2001; Bowman, Koller & Hodges, 1997) where the user can only travel in the direction they are looking. Body-direction travel, where the heading of the body and the direction of travel are locked together, but the direction of view can be manipulated independently. Lastly, independent movement, where travel is independent of both the viewing direction and body direction.

A characteristic of VE navigation is that people tend to travel in paths that are generally straight (Ruddle & Jones, 2001). This is perhaps caused by the design of the movement interface, and if a greater number of DOFs are available to a user (e.g., by implementing independent rather than view-direction travel) then it will be easier for them to deviate from a straight-line path. Of course, people are more likely to exploit these additional DOFs if they are controlled in a coordinated and, ideally, natural manner. However, the greater the number of DOFs available to the user the greater the cognitive effort required to control the interface, so a high DOF interface generally takes more time to learn than one with fewer DOFs. Thus, interfaces that provide a large number of DOFs (flexible movement) have both advantages and disadvantages.

One of the main ways of displaying immersive VEs is via a HMD, the position and orientation (pitch, yaw, and roll) of which can be tracked using a variety of the tracking devices outlined in Chapter 1. These technologies allow rotation and position change in the real world to be used to update rotation and position in the VE. The HMD uses the same muscle groups (i.e. neck and leg muscles) as those used for looking around and travelling through a real environment and therefore provides a more familiar way to vary one's view of a VE than using a mouse and keyboard, which are the devices most commonly used for navigating desktop VEs.

Depending on the tracking devices that are used, an HMD may be used in either a "tethered" mode or a "roving" mode. With a tethered HMD the user's physical changes of orientation are tracked, but changes of position are controlled using an abstract interface such as holding down a button on a 3D mouse. With a roving HMD the user's physical position and orientation are tracked, so the user may walk through the environment as if they were physically there. However, a roving HMD can only be used if a large empty physical space is available, and the cost of extended range position tracking can be justified.

Research performed using a variety of basic spatial tasks indicates that physical rotational movements are much more important than translational movements. Direction estimates between objects in a room-sized space were made significantly more accurately when participants physically rotated than when they imagined they had done so (Presson & Montello, 1994; Rieser, 1989). Path integration is accurately performed if participants actively or passively make physical turns, but not if they make virtual turns (visual information presented on an HMD) or observe someone else physically walking (Klatzky, Loomis, Beall, Chance & Golledge, 1998). Judgements of angles turned though are more accurate when participants physically turn, with or without visual information, and less accurate if only visual information and no body-based information is provided (Bakker, Werkhoven & Passenier, 1999). However, in a comparable, and more recent experiment (Expt. 1, turn&go) by Riecke, van Veen, and Bülthoff (2002), participants viewed a VE, a field of blobs, on a 180 degrees FOV projection screen and navigated via a three button interface. Contrary to the errors found in the study by Bakker et al. (1999) the optic flow information provided by the field of blobs seems to have been sufficient for accurate visual translations and rotations. Riecke et al. suggested that the projection screen might have provided an adequate frame of reference for the task. They also conducted two paper-and-pencil spatial ability tests and found that they correlated with participants' navigation ability in two other, similar, experiments, suggesting that spatial ability is also a contributing factor to performance.

Studies have also investigated the contribution of vestibular and proprioceptive cues for complex spatial tasks and acquiring knowledge about large-scale spaces. Waller, Loomis, and Steck (2003) compared participants' ability to point and estimate distances between five places in a previously unknown real-world environment. The first group (full-cue) were driven along a 1600m round trip while a camera, fixed inside the car, recorded the journey. Another group (vestibular) sat in the back and watched the scene through the camera via a HMD, they therefore received the same vestibular cues as the full-cue group. The third group (non-matching) watched the video of the journey via an HMD while they were driven on a different trip, and therefore received conflicting vestibular cues to the full-cue and vestibular groups. The last group (video) watched the video on their HMD while seated stationary in a laboratory. There were only significant differences between the full-cue condition and the three HMD conditions. The mismatch or even the absence of vestibular cues made no significant difference to the acquisition of spatial layout between the three HMD conditions. Waller et al. (2003) suggested that factors such as the FOV, visual fidelity, and the ability to actively look around affect the acquisition of spatial knowledge of larger scale spaces more than vestibular cues. However, this is contradicted by a subsequent study by the same group of researchers. Waller, Loomis, and Haun (2004) investigated the spatial knowledge gained from an 840 metre journey through a university campus. One condition physically walked along the route and viewed the scene via a camera that simultaneously displayed the images to the walking participants' HMD and recorded the images for later use. The second group watched the previously made video and their direction estimates between the five locations were significantly less accurate than the group that had learned the route by walking, although the magnitude of the difference was relatively small.

The provision of vestibular and proprioceptive cues just for rotational movements generally has little effect when participants are asked to perform search tasks in virtual spaces. In a similar task to the one used throughout this thesis, Ruddle and Jones (2001) investigated participants' ability to search a VE using either a desktop monitor or a tethered HMD. They found that there were no significant differences between the two conditions; in both cases participants had great difficulty completing the task in a fifth of the trials.

Ruddle, Rayne, and Jones (1999) compared the effects of two display devices (tethered HMD and a desktop monitor) on participants' ability to learn the layout of a complex large-scale VE with more than 70 rooms. Participants that used the HMD navigated the buildings significantly quicker and were significantly more accurate in estimating relative straight-line distances between previously seen targets than those that used the desktop monitor. However, more recent research using a similar task found that participants estimated relative straight-line distances more accurately with a desktop display than a tethered HMD (Ruddle & Péruch, 2004).

The provision of vestibular and proprioceptive feedback for rotational and translational movements in a VE is beneficial for the navigation of large-scale spaces. Chance, Gaunet, Beall and Loomis, (1998) asked participants to travel through a 3x4 metre virtual maze, displayed via a HMD, and point to objects that had been encountered en-route. There were three experimental groups: participants either physically walked though the maze, made physical rotations to look around while translating with a joystick, or performed both rotations and translations with the joystick. Physical walking produced significantly more accurate direction estimates than those that used the joystick for movement.

In another complex task (Grant & Magee, 1998), participants were asked to estimate the direction of the start position at eight locations along a route though 7400 square metres of a real museum or low visual fidelity virtual museum. In the training session participants either physically walked though the real museum (RW), or viewed the virtual museum via a rotationally tracked HMD where they either walked in place (VE-walk) or used a joystick to travel (VE-joystick). Results showed that participants that physically walked through the museum produced significantly more accurate direction estimates than those in either of the two VE conditions that travelled whilst stationary. Participants were then asked to visit the eight targets in a new sequence where the shortest route was not the route participants were asked to take in the training session. Participants who used the joystick, which provided minimal proprioceptive feedback, travelled a significantly greater distance than those in the other two conditions. Proprioceptive feedback, provided by physical walking or by the VE-walk movement technique, a pawing foot motion, allowed participants to gain greater knowledge of the spatial layout of the museum so that they were able to take short cuts to the next target.

The possible benefits of walking in place over more abstract movement techniques have also been investigated in other studies. Slater, Usoh, and Steed (1995) asked two groups of participants to pick up an object and move it to another virtual room. Participants that used a walking in place technique reported an increased sense of presence over those that indicated their desired direction by pointing and translating with the use of a mouse. However, participants preferred to use the pointing technique and there were no metrics used with which to compare task performance. More sophisticated algorithms for walking in place have been developed by Templeman, Denbrook and Sibert (1999).

Finally, Zanbaka, Lok, Babu, Xiao, Ulinski and Hodges (2004) compared participants' ability to gain spatial knowledge from a VE using one of four methods of travel that ranged from navigating while stationary (they looked at a monitor and travelled using a joystick) to physically walking through the VE (real walking with an HMD). Participants that walked in the VE gained more knowledge and understanding of the VE than participants that used the other methods of travel. This suggests that the more familiar the method of travel the more cognitive resources can be allocated to understanding the scene.

To summarise, all real-world human navigation involves some form of physical motion. Information from translations and rotations of locomotion are combined with visual information, and other sensory information, to aid efficient navigation however, little is known about the contribution made by each source. Participants frequently encounter difficulties navigating VEs when they are physically stationary and when only visual information is presented, for example, on a desktop monitor (Ruddle & Jones, 2001). Combining physical rotation with an abstract means to changes one's position allows simple navigational tasks such as path integration to be accurately performed (Presson & Montello, 1994; Klatzky et al., 1998) but do not seem to aid more complex tasks such as learning the layout of a virtual building (Ruddle et al. 1999; Ruddle & Péruch, 2004). However, walking through an environment and changing both position and orientation in a physical, rather than abstract, manner produces a modest but significant improvement in performance when learning the location of objects along a route, (Chance et al., 1998; Grant & Magee, 1998; Waller et al. 2004).

Although studies have shown that it is possible to learn complex navigation tasks in desktop VEs (Ruddle et al. 1997), spatial learning takes place substantially more slowly and is less accurate when compared to the knowledge gained from the real world. Both visual content of the scene and the FOV have been shown to effect participants' ability to navigate efficiently. Increasing the visual content of the scene increases the amount of landmark information available, allowing near perfect performance for simple spatial tasks (Riecke et al. 2002) but not more complex tasks (Witmer et al. 1996). A restricted FOV has been shown to effect the perceptual layout of spaces, but only when it is 22 degrees or less (Alfano & Michel, 1990). Although judgements of ego rotations are performed more accurately with a wide FOV than with a narrow FOV (180 degrees vs. 24 degrees), with a wide FOV errors quickly increase as a task becomes more cognitively demanding (Riecke et al., 2002; Bakker, et al., 1999).

2.4 Metrics

Ruddle and Lessels (2004) categorise three levels of data that can be used for the evaluation of navigation. The top level metric (level 1) measures task performance. The middle level metric (level 2) relates to the physical behaviour of participants and examines the actions they make in order to achieve their navigation goals. The lowest metric (level 3) examines the rationale of participants, and attempts to uncover the decision making process that led to the participants adopting certain behaviours. The majority of studies that investigate navigation in VEs use level 1 metrics to evaluate aspects of the VE system, with common examples being time taken, distance travelled, and the number of correct/incorrect turns made.

Level 2 metrics attempts to provide an explanation as to how performance was achieved, examples include the examination of movement subtasks (e.g. time spent moving vs. looking around), and how the interface was used (e.g. proportion of time spent using various controls). It is important to analyse users' navigational behaviour, because subtle changes in that behaviour can lead to drastic changes in navigational performance. In other words, it is behaviour that often dictates performance. Relatively few present studies have analysed behavioural metrics, however, one recent study that has is by Czerwinski et al. (2002). Participants were asked to find target objects and move them to predefined locations as quickly as possible. With a novel navigation technique (the faster participants moved forward the higher and steeper they viewed the VE) males performed faster then females with a narrow FOV, but with a wide FOV females' speed increased to equal that of males. Data analysis was not only performed on performance results (speed) but also on how participants achieved those results. The behavioural metric showed that males travelled higher than females with the narrow FOV, while both genders travelled higher with the wide FOV. Also, in the wide FOV, there were differences in the use of strategy between the genders, females travelled less distance to achieve a faster time, while males travelled further, but quicker, to achieve a faster time.

Another study to investigate movement behaviour, in this case search strategy, was Darken and Sibert (1996). Participants were asked to find five ships in a large-scale VE in four conditions that varied in the amount of navigational assistance. Participants that were provided with a map searched using a systematic lawnmower strategy, and when a radial grid was superimposed over the environment or map they used the lines as paths to structure their search. Participants with no map or grid assistance, however, followed the coastline of the islands and executed ineffective searches, often searching the same place more than once, travelled further and took more time to complete the task. These behavioural metrics show the different strategies adopted and illuminate how performance results occurred.

Level 3 metrics refer to participants' rationale for making decisions while navigating. Techniques include recording on and off screen activates and conversations during and/or after navigation (Murray, Bowers, West, Pettifer & Gibson, 2000; Gamberini, Cottone, Spagnolli, Varotto & Mantovani, 2003) or asking participants to 'think aloud' to verbalise their thoughts whilst engaged in the navigation task (Darken & Banker, 1998; Sullivan, Darken & McLean, 1998). However, while these metrics give valuable insight into the navigation process, they have only occasionally been used in studies of navigation and were not used at all during the studies described in this thesis.

2.5 Summary

The navigation research reviewed in this chapter highlights many gaps in our understanding of the way FOV, visual fidelity, and the movement interface affect navigation in virtual environments. The experiments in this thesis investigate the effects of a restricted FOV in a real and virtual environment; the affects of three levels of visual scene fidelity; and the affects of the movement interface and physical movement, during a spatial search task in virtual spaces. The results of the experiments are analysed using two of three levels of metric (see section 2.4, i.e. performance and behaviour). Moreover, interactions between these three aspects of a VE application are also studied and analysed to provide a more comprehensive understanding of how humans navigate virtual spaces.

Chapter 3

Overview of the Experiments

3.1 Introduction

As the general format of all the experiments was similar, and to avoid repetition, this chapter provides an overview of the experimental task, method, procedures, and conditions which are used throughout the studies. The search task used throughout all of the experiments reported in this thesis is based on the task originally devised by Ruddle and Jones (2001). In that original study participants were asked to travel around a small 10 metre x 10 metre cluttered environment searching for eight targets in amongst 16 possible locations. To do this search task efficiently participants had to remember where they had travelled and minimize the extent to which they retraced their steps.

3.2 Materials

In the original study the presence of a blue-topped cylinder signalled a possible target location (see Figure 3.1a), and targets were placed in a recess in the blue top so they were visible whenever participants were within a distance of 0.747 meters and looking in the appropriate direction. However, during data analysis it was not possible to distinguish between occasions when participants travelled to a blue-topped cylinder to check for a target inside, and occasions when participants passed one of the cylinders while en-route to another.

To help analyse participants' behaviour as they conducted the search task in the present experiments this ambiguity was prevented by a small modification to the design of the cylinders. Targets and decoys were indicated by the presence, on top

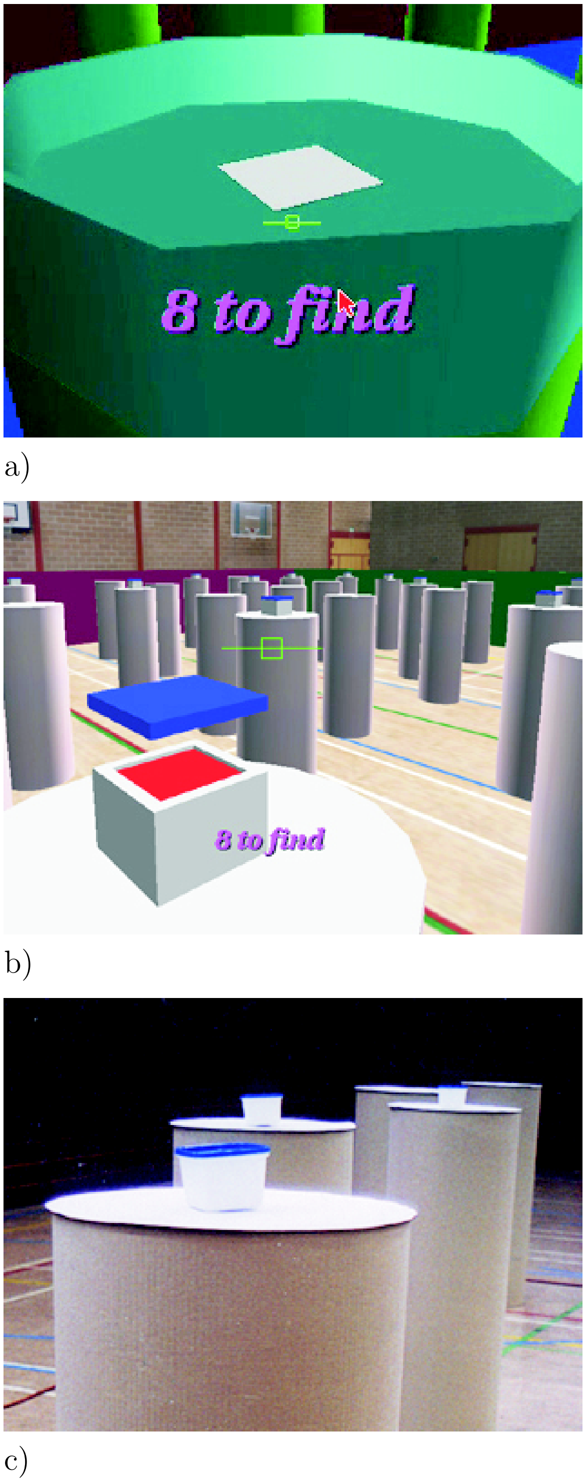


Figure 3.1: a) The target and decoy cylinders used in the original study of Ruddle and Jones (2001). b) The target and decoy boxes used in the sports hall VE studies. c) The white plastic boxes used in the real world study.

of a cylinder, of a small white box with a blue lid. Target boxes each contained a target object (a red square; see Figure 3.1b) while the decoy boxes were identical but with the red square absent. In the real world study the box and lid were plastic and the target object was a piece of red card. With this design if a participant wanted to search a box for a target they had to raise the lid and look inside. This indicated that the participant was making a conscious decision to search for a target in that box.

The VE application used for all of the virtual experiments was a modified version of the software used in the original Ruddle and Jones (2001) study, adapted to reflect the changes made for data analysis to the decoy and target boxes. The VE software was written in C++ and OpenGL Performer and ran on a SGI Onyx 3400 and on a SGI Onyx 4.

3.2.1 Environments

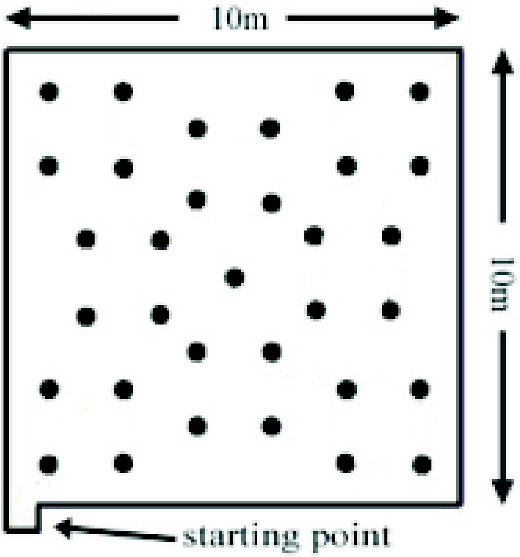


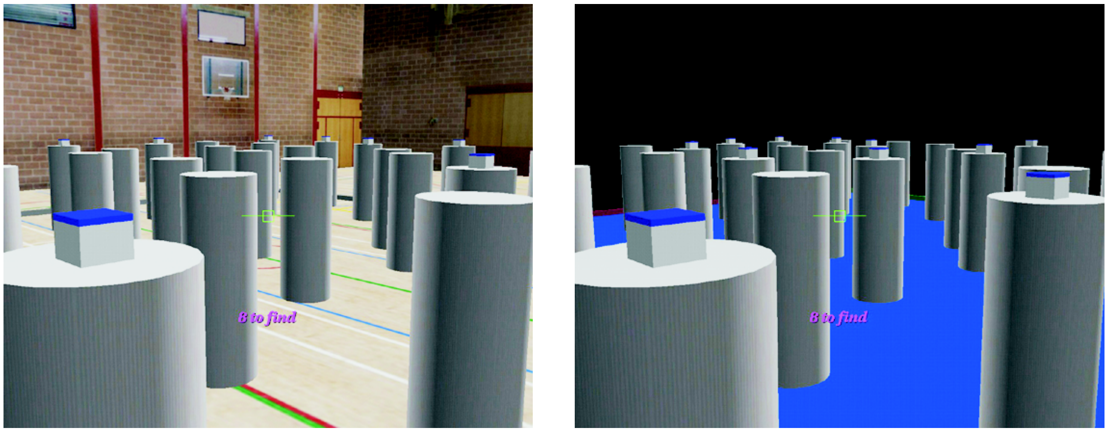
Figure 3.2: The configuration of the cyclinders, eight groups of four with the 33rd in the centre.

The experimental environments all had the same general structure. This comprised 33 cylinders, 32 of which were arranged in eight identical groups of four, with the 33rd cylinder positioned in the centre (see Figure 3.2). All cylinders measured 1.35 metres high and were either 0.5 or 0.15 metres in diameter, depending on the particular environment. The positions of the target and decoy boxes, for each trial, were defined by the same rule used to position the blue-topped cylinders in the original study. That is, within each group of four cylinders one was randomly chosen to be the target and another a decoy. This ensured that the targets and decoys were distributed around the environment.

Figure 3.3: Photograph of the real world environment.

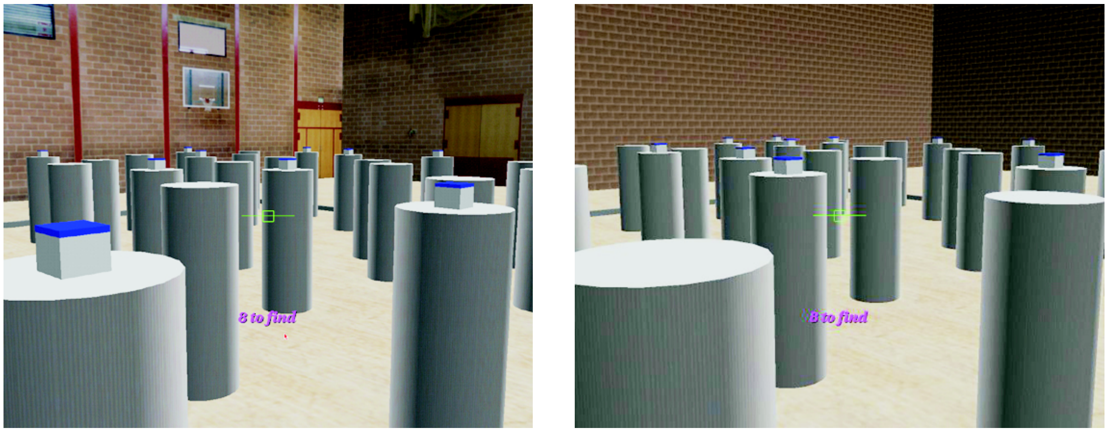
One real-world and six virtual environments were created. The real world environment location was the University of Leeds sports hall (see Figure 3.3). A total of 36 cylinders were made from corrugated paper, 33 cylinders were arranged inside the sports hall while three were placed outside and used to explain the task to participants. A climber's rope was placed around the perimeter of the cluttered environment, substituting the coloured walls that defined the space in the original VE. On top of 16 of the cylinders was placed a small plastic box, with a blue lid. A target object was a small square piece of red card inside a box. Participants carried with them eight blue cards and deposited them on top of the eight red cards, this replicated the on-screen counter and the target objects that turned from red to blue in the original study.

The high fidelity sports hall VE (see Figure 3.4a) was a photo realistic model of the real world sports hall and used textures captured by digital camera from the floor, walls, and ceiling of the real sports hall. The design of the target and decoy



a) b)

Figure 3.4: a) High fidelity sports hall VE. b) Low fidelity sports hall VE.



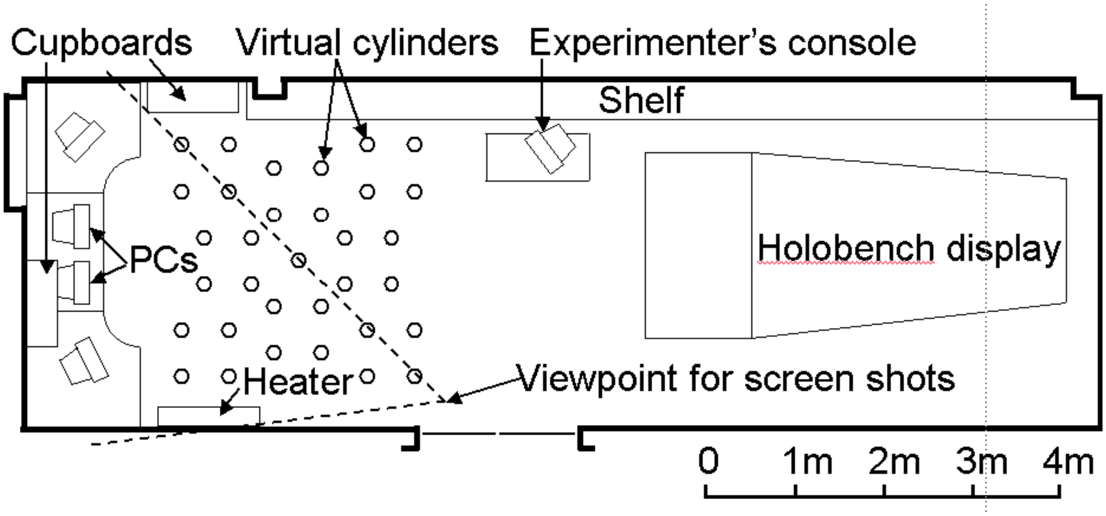
a) b)

Figure 3.5: a) High fidelity sports hall VE with no floor lines. b) Medium fidelity (brick tiles) sports hall VE with no floor lines.

boxes replicated the plastic boxes used in the real world study, and a 0.1 metre high perimeter grey curb defined the navigable space. The low fidelity sports hall VE (see Figure 3.4b) was the same as the high fidelity sports hall VE except that the sports hall textures and geometry that surrounded the navigable space was absent, each low perimeter curb was a different colour, and only the navigable space had a surface which was coloured blue. This VE most closely resembled the VE used in the original Ruddle and Jones (2001) study.

Two other virtual models of the sports hall were created. One was identical to the high fidelity sports hall except that the floor lines used for sporting activities (basketball and badminton etc.) were absent (see Figure 3.5a). The other used the same geometry and floor texture as the line-less high fidelity sports hall VE but the walls were rendered using a small tiled brick texture (see Figure 3.5b).

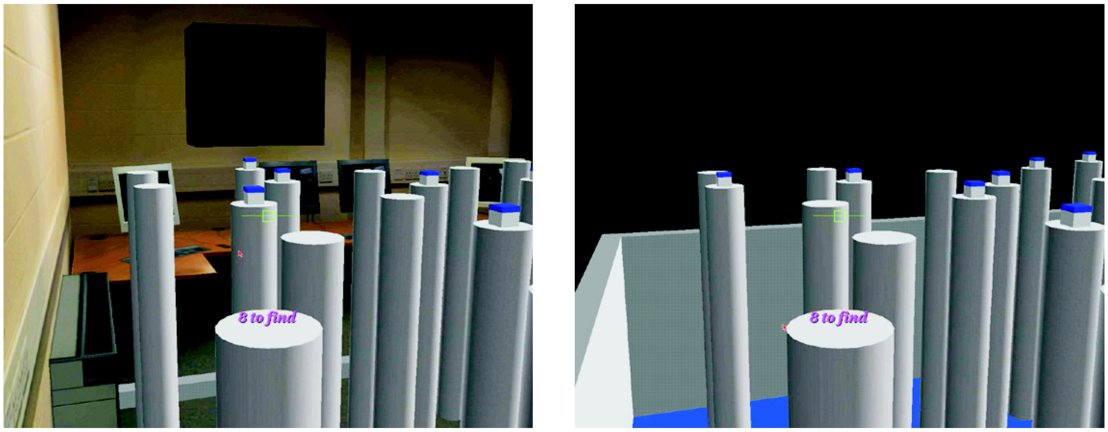
a)



b)

Figure 3.6: a) Photograph of the University of Leeds School of Computing virtual reality laboratory. b) Plan view of the laboratory.

The fifth VE was a photorealistic model of the School of Computing virtual reality laboratory (see Figure 3.6) and was constructed using images of the laboratory surfaces as textures which, like the high fidelity sports hall VE, were captured using a digital camera and used to create 38 separate textures ranging from 4 Kb to 5.8 Mb and totalled 30 Mb in RGB format. The cylinders in this environment were 0.15 metres diameter, to allow all 33 to fit within the space constraints of the lab (see Figure 3.7a). The 6th VE had the same cylinder layout as the photorealistic VR lab, but without the textures (see Figure 3.7b).



a) b)

Figure 3.7: a) The high fidelity virtual reality laboratory. b) The low fidelity virtual reality laboratory.

3.2.2 Display and Field of View

The real world experiment had two FOV conditions, full-view and restricted-FOV. There was no restrictions to participants in the full-FOV condition, but participants in the restricted-FOV condition wore modified safety goggles (see Figure 3.8) that reduced each eye's FOV down to approximately 20 degrees on the horizontal and 16 degrees in the vertical direction.

In the normal-FOV desktop conditions the VE was viewed with a 48 x 39 degree graphical FOV via a CRT colour monitor with a 475 mm x 300 mm viewable screen size. The resolution was 1280 x 1024 pixels and the refresh rate was 72 Hz. All participants viewed the monocular displayed VE from a distance of approximately 60 cm. Participants in the wide-FOV conditions used the same monitor as the normal-FOV conditions but with two additional monitors of the same type and size

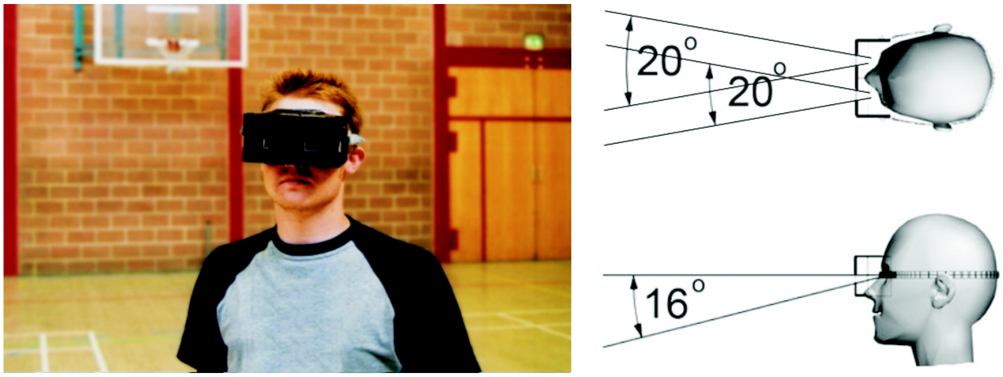


Figure 3.8: Goggles used in the real world restricted-FOV condition.

placed on its left and right. Each monitor displayed 1 of 3 viewing frustums at a resolution of 800 x 600 pixels creating a 144 degrees horizontal graphical FOV.

Both the HMD conditions used a Virtual Research V8 HMD (see Figure 1.9) that displayed two stereo images separated by a participant's interpupillary distance (IPD). Each VE was viewed with a GFOV of 48 degrees horizontal x 36 degrees vertical, which matched the PFOV of the HMD. Each LCD screen provided a 640 x 480 pixel display at 60 Hz.

3.2.3 Movement Interface

There were two desktop movement interfaces, providing forward-only and 4-way movement control. Participants travelled around the VE using the keyboard cursor keys for movement across the horizontal plane. Motion was continuous and at a speed of either 1 m/s (sports hall VEs) or 0.5m/s (VR lab VEs) while the cursor keys were depressed, and stopped when released. In the forward-only condition (view-direction movement) participants could only travel forward in the direction in which they were looking, achieved by holding down the 'up' cursor key. In the 4-way (independent movement) condition they could move forward, back, left, and right across the horizontal plane of the VE and could also travel diagonally by holding down pairs of keys (e.g. forward and left). In both of these movement interfaces the mouse was used to control the participant's direction of view. Looking up and down was achieved using zero order control; by moving the cursor up the screen the viewing pitch increased by up to +90 degrees, and the viewing pitch decreased by a corresponding amount if the cursor was moved down the screen. Looking left and right was accomplished using first order control with the rate of turning increasing proportionally with the cursor's distance away from the vertical centreline of the screen. The maximum rate of turning was 135 degrees/second.

Participants could raise and lower the lids of the target and decoy boxes by pressing the left mouse button. Pressing it once raised the lid, and pressing it again lowered the lid. If there was a target present inside the box the participant pressed the right mouse button to select it, the target then turned from red to blue indicating that it had been found, comparable to the depositing of a blue card on top of a red target in the real world experiment. The lid was then lowered automatically by the VE software. The software prevented participants from moving away from any box until its lid was lowered. A lid could only be raised, and a target selected, if the participant was within 0.747 metres of the centre of the box (i.e. they were adjacent to it), it was within the participant's FOV, and the box's cylinder was the closest cylinder to the participant's virtual self.

There were also two HMD interfaces, tethered and roving. Both conditions used an HMD (see Figure 1.9) with an Intersense Inertiacube2 inertial sensor attached to track the participant's head pitch, roll, and heading (yaw) which were then used to update the view of the VE. In the tethered HMD condition participants controlled their forward movement by holding down a button on a 3D mouse (see Figure 1.5), the speed of movement was the same as for the desktop interface. In the roving HMD condition, however, the participant physically walked around the laboratory and their position was tracked by a World Viz Precision Position Tracker and used to update the participant's position in the VE.

3.3 Experimental Conditions

A total of four behavioural experiments were performed, summarised in table 3.1. In the real world experiment there were two FOV conditions, full-FOV and restricted-FOV. Particpants in the full-FOV condition conducted the search task with a full human FOV, while the restricted-FOV participants searched the space whilst wearing view restricting goggles.

Experiment 2 and 3 used VE models based on the sports hall. Experiment 2 contained four conditions, and participants searched the VE using one of two movement interfaces, forward-only or 4-way movement control, and one of two visual scenes, high or low fidelity. Experiment 3 also contained four conditions; participants searched using either a normal FOV, via one monitor, or a wide FOV, via three monitors, while viewing either a high or medium fidelity visual scene. The high fidelity scene was essentially the same as the VE used in the second experiment

Table 3.1: The conditions and procedures used in the four experiments.

Experiment	**Environment**	**Interface**	**Display**	**FOV (degrees)**	**number of practice trials**	**test trials**
1	real-world sports hall	real walking	real world	200 x 120	1 (real world)	3
	real-world sports hall	real walking	real world	20 x 16	1 (real world)	3
2	low-fi sports hall VE	forward-only	desktop monitor	48 x 39	2 (desktop)	4
	hi-fi sports hall VE	forward-only	desktop monitor	48 x 39	2 (desktop)	4
	low-fi sports hall VE	4-way movement	desktop monitor	48 x 39	2 (desktop)	4
	hi-fi sports hall VE	4-way movement	desktop monitor	48 x 39	2 (desktop)	4
3	brick-tiled sports hall VE with no floor lines	forward-only	desktop monitor	48 x 39	2 (desktop)	4
	hi-fi sports hall VE with no floor lines	forward-only	desktop monitor	48 x 39	2 (desktop)	4
	brick-tiled sports hall VE with no floor lines	forward-only	desktop monitor	144 x 39	2 (desktop)	4
	hi-fi sports hall VE with no floor lines	forward-only	desktop monitor	144 x 39	2 (desktop)	4
4	hi-fi VR laboratory	forward-only	desktop monitor	48 x 39	4 (desktop)	4
	hi-fi VR laboratory	tethered HMD	HMD	48 x 36	2 (desktop) + 2 (HMD)	4
	hi-fi VR laboratory	roving HMD	HMD	48 x 36	2 (desktop) + 2 (HMD)	4
	low-fi VR laboratory	forward-only	desktop monitor	48 x 39	4(desktop)	4
	low-fi VR laboratory	roving HMD	HMD	48 x 36	2 (desktop) + 2 (HMD)	4

except that the floor consisted of a repeated base texture, as the sports court lines were absent. The medium fidelity VE contained less detail on the walls because they were rendered with a small brick texture.

Experiment 4 used VE models based on the School of Computing virtual reality laboratory. The main experiment contained three conditions, desktop, tethered HMD, and roving HMD, and each used a high fidelity visual scene. The supplementary experiment contained two conditions, desktop and roving HMD, and used low fidelity VEs. Participants in the two desktop conditions searched the VE while seated stationary in a chair and viewed the VE via a desktop monitor. In the three HMD conditions participants viewed the VE via a HMD that was tracked in three axes of rotation. In the tethered HMD condition participants stood and turned in place to view the VE while travel was achieved by holding down a button on the 3D mouse. The two roving HMD conditions were similar except that participants physically walked to update their position in the VE. All three HMD conditions were conducted in the real VR laboratory while the two desktop conditions were conducted in another room.

The motivation for these various experiments was as follows. The real world experiment (Experiment 1) was conducted to investigate two factors. First to confirm what had been assumed in the previous VE study (Ruddle & Jones, 2001) that this task would be trivial to perform in the real world, and to provide a “gold standard” of participants’ performance for future research. Second, to investigate participants’ performance when their FOV of the real world scene was substantially restricted. Experiment 2 was then conducted to bridge the gap between the original VE study and the real world environment used in Experiment 1 by investigating the effect of two different desktop movement interfaces and two visual scene characteristics on participants’ search performance in a desktop VE. Experiment 3 was conducted to further investigate the possible effects of different FOVs and two visual scene characteristics on participants’ search performance, again in a desktop VE. Finally, Experiment 4 was conducted to investigate three different movement interfaces that differed in terms of the proprioceptive and vestibular information that was provided for rotational and translational movements.

3.4 Procedures

The procedures for the experiments followed a similar procedure to that of the original Ruddle and Jones (2001). Each participant first practiced using the interface until they could fluently use the controls. The participant then performed practice trials that allowed him or her to become familiar with the search task, and then completed some test trials. All participants were asked to minimize their journey path and to avoid checking each possible target location more than once. Each trial began at the stating point in the boundary recess (see Figure 3.2) and participants searched until they had found all eight targets. Participants were informed that the targets were always in the white boxes, but that their positions changed between trials. No feedback was provided on participants' performance or their search strategy. The exact details of the interface and the number of practice and test trials differed between the sets of experiments (see table 3.1).

Chapter 4

Real World Navigation and FOV

4.1 Introduction

The experimental task for the real world experiment was essentially the same as the task used for the original VE study by Ruddle and Jones (2001) and was conducted to investigate two factors. First to confirm what had been assumed in the previous VE study, that this task would be trivial to perform in the real world, and to provide a "gold standard" of participants' performance for future research. Second, to investigate participants' performance when their FOV of the real world scene was substantially restricted (to 20 x 16 degrees in each eye). Considering that a restricted FOV has been shown to hinder participants' ability to perform simple real world spatial tasks, it was hypothesized that restricting the FOV for this real world search task would produce a decrease in performance (slower search or greater distance travelled) when compared to participants with no visual restrictions.

The experiment was conducted over two days in the university's sports hall. A between participants design was used, with participants randomly allocated to either the *full-view* (normal) or *restricted-FOV* condition.

4.2 Method

4.2.1 Participants

Ten participants took part in the experiment. Their ages ranged from 18 to 36. All participants were either graduates or undergraduates who volunteered for the experiment and were paid an honorarium for their participation. All gave informed consent.

4.2.2 Materials

The real world study was held in the university sports hall (Figure 3.3) and the design of the environment was comparable to the original VE study, see Chapter 3 for a description. In the restricted-FOV condition, participants wore modified safety goggles (see Figure 3.8) that restricted the participants FOV, see Chapter 3 for a description.

4.2.3 Procedure

The procedure for the real world experiment was similar to the procedure used in the original VE study. Each participant performed four trials. The first was treated as a practice trial and the three subsequent trials were treated as test trials. Participants performed the trials individually and took approximately 35 minutes to complete the experiment.

For each trial, a participant was given eight pieces of blue card and asked to walk around the environment depositing these cards on top of the eight red cards (target objects) inside the target boxes. This ensured that if they revisited a target box during a trial, they would know this by the presence of the previously deposited blue card. Participants were asked to walk at a normal speed, minimize their journey path and avoid researching possible target locations They were also asked to place each box lid back the way they had found it so that they wouldn't be able to know, simply by looking at the lid, if they had already visited that box during that trial. To prevent participants seeing the positions of the boxes before the beginning of the trial, they waited outside the sports hall while the targets and decoys were placed in position, and then blindfolded while being guided to the starting point for each trial (the boundary recess; see Figure 3.2). The start of the trial was signalled by the removal of the blindfold.

Participants searched until they had found all eight targets, and then left the hall and waited outside while the boxes were repositioned ready for the next trial. No feedback was provided on participants' performance or the search strategy that had been adopted. The procedure for the participants under the restricted-FOV condition was the same except they conducted the trials whilst wearing the view restricting goggles. These participants were asked to look at the floor while they were guided to the starting point, and the restricted view rendered the blindfold unnecessary.

The positions of the target and decoy boxes, for each trial, were defined by the same rule used to position the blue-topped cylinders in the original study. That is, within each group of four cylinders one was randomly chosen to be the target and another a decoy. This ensured that the targets and decoys were distributed around the environment.

During each trial three types of data were recorded. First, the time that each participant took to complete the task was recorded. Second, the route that each participant travelled was sketched on a plan of the environment. Finally, each trial was recorded on videotape, which was then used after the experiment to confirm both the time taken and the route travelled.

4.3 Results

Participants' performance in each trial was measured using two primary types of data:

1. Task performance
    (a) Time taken to find the eight targets
    (b) Total number of visits to target and decoy boxes during a trial
    (c) Distance travelled (percentage above the optimum route length)
2. Behaviour
    (a) Search strategy

Statistical analyses of the data were performed using mixed design analyses of variance (ANOVAs) that treated the field of view as a between participants factor (full-FOV vs. restricted-FOV) and the trial number as a repeated measure. Only data for the three test trials were analysed. None of the interactions were significant.

4.3.1 Task performance

Due to an error, time data for the first test trial of one participant were not recorded. The time that participants took in the full-FOV (M = 94.4 s, SD = 26.9) and restricted-FOV conditions (M = 104.3 s, SD = 22.0) was similar ($F(1, 7) = 0.95$, $p > .05$). Also, the time taken in the test trials did not change significantly as the test trials progressed ($F(2, 14) = 0.47$, $p > .05$). Participants started to search the environment as soon as the trail started, stopped momentarily (or sometimes just slowed down) to check the boxes, and rarely stopped between target or decoy boxes to look around.

The total number of visits to target and decoy boxes did not differ between FOV conditions (M = 15.3 in both cases), or between the three test trials. In 27 of the 30 trials participants did not revisit any target or decoy boxes, meaning that these trials were completed with optimal efficiency according to this metric. In the full-FOV condition, one participant revisited two boxes in one trial (see Figure 4.1) and, in the restricted-FOV condition, two participants each revisited one box in one trial (see Figure 4.2). In the full-FOV condition participants found all eight targets before reptetition in 93% of trials, while in the restricted-FOV condition 86.7% of trials were conducted perfectly.

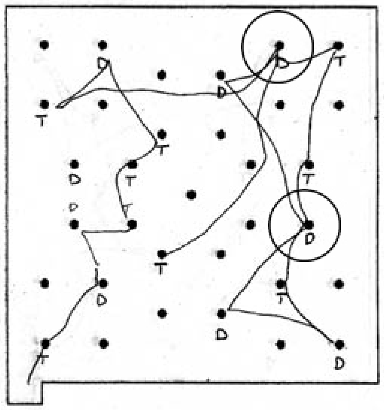


Figure 4.1: The path shows that the participant in the full-FOV condition revisited two boxes (circled). T = target, D = decoy.

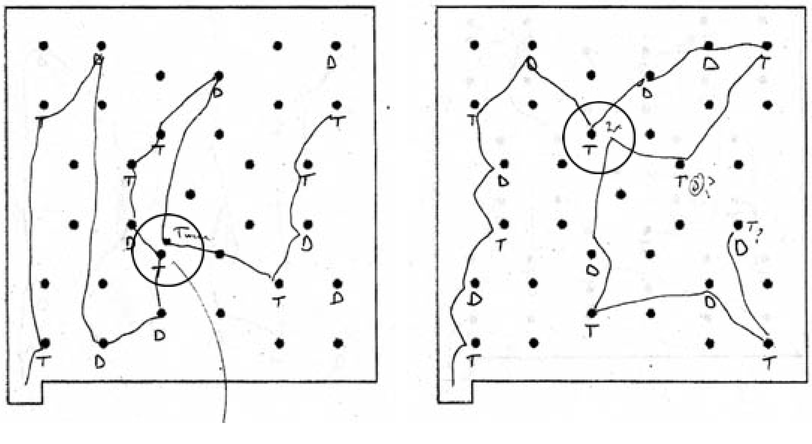


Figure 4.2: Two paths walked by two participants who revisted one box each (circled), in the restricted-FOV condition.

The distance that participants travelled was compared to the shortest possible distance, calculated using a travelling salesperson problem (TSP) algorithm. First the distance travelled by a participant in a trial was approximated by calculating the straight line distance between the start point and the centres of the target and decoy boxes, in the order that they were visited.

The TSP program used to calculate the shortest possible route was written in C++ by the authors and utilized an algorithm obtained from the Combinatorial Object Server (Ruskey & Sawada, 2002). Unlike conventional TSP algorithms, the software implemented did not have to find a solution that started and finished in the same place. Instead it found the shortest route for a one-way, outward-bound trip, which ended at the last visited target.

An example of an actual route taken by a participant in a trial and the corresponding solution calculated by the TSP program is shown in Figure 4.3. In this example the participant did not visit the last decoy box because the task was complete when the eighth, and final, target was found. The program, however, included this decoy (circled) as part of its initial solution but then remedied the inconsistency by subtracting the last route segment from the distance that was calculated. The TSP program then drew the shortest route on a plan view using OpenGL.

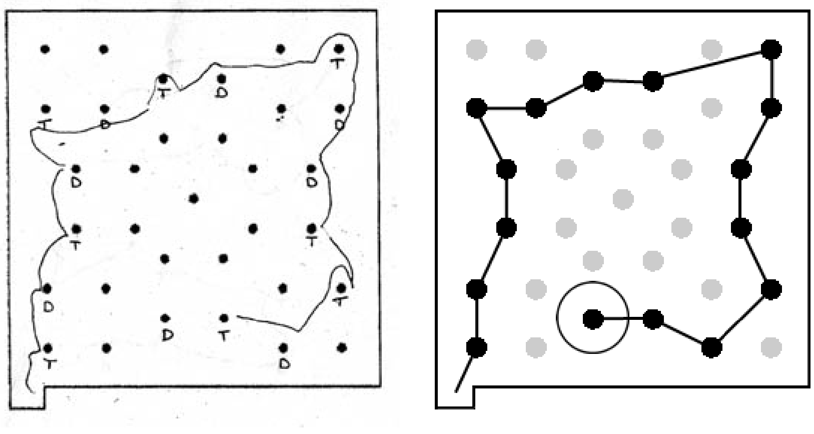


Figure 4.3: A sketch of a participant's route through the environment in the real world experiment (left), T = target, D = decoy. The shortest route, as calculated by the TSP program (right). The route ends at a decoy (circled) so the program subtracted the distance of the last route segment. In this trial, the participant followed the shortest route.

In each trial, the distance that participants travelled was derived by expressing the distance that participants travelled as a percentage above (or below) the distance of the shortest possible route. There was no significant difference in this percentage between full-FOV (M = 10.5%, SD = 13.5) and restricted-FOV conditions (M =15.0%, SD = 18.4), ($F(1, 8) = 0.37$, $p > .05$), or between the three test trails ($F(2, 16) = 0.03$, $p > .05$). The only trial in which there was an exact match between the path taken by a participant and that calculated by the TSP program is the one shown in Figure 4.3.

Some participants walked a path shorter than the TSP program solution because, while searching for the eight targets, they passed some decoy boxes and fortuitously left them un-searched. A similar behaviour was also observed in the original VE study where participants travelled past a decoy but did not search it.

Considering these performance measures, as predicted, participants completed the task in the full-FOV condition with near perfect efficiency. It was hypothesised that wearing view restricting goggles would reduce participants performance, however, even with a restricted FOV the task was trivial to perform.

4.3.2 Behaviour

Inspection of the travelled paths showed that participants usually started their search by either following the perimeter of the VE or adopting a lawnmower-type pattern. In most trials, participants found the majority of the targets using one of these two strategies. Any remaining targets were then searched for using secondary strategies, examples of which included spiralling in on the centre of the VE after completing a search of the perimeter, and the somewhat random searches that occurred when participants were unsure of which targets and decoys they had already visited.

Interest in the present study is centred on participants' initial (primary) search strategies. For each trial, these were analyzed using a three-stage process:

1. Classifying the strategy as *perimeter*, *lawnmower*, or *other*.
2. Counting the number of passes made during the search.
3. Counting the number of targets found before any revisitation.

Searches were classified by dividing the VE into four quadrants and noting the order in which these were visited. Perimeter searches visited the quadrants in the order 1-2-3-4-1 (clockwise search; see Figure 4.4 a) or 1-4-3-2-1 (anticlockwise). Lawnmower searches involved a sequence of passes that crossed the VE's centreline, progressing along the centreline from one side of the VE to the other. The centreline was always perpendicular to the direction of the passes, so in some trials this was the dividing line between quadrants 1/2, and 3/4 (see Figure 4.4 b), but in other trials it divided quadrants 1/4 from quadrants 2/3. All lawnmower searches were predominately in line with the circulation routes created by the structure of the cylinders; no lawnmower search was conducted that progressed diagonally across the environment. One search in each FOV condition could not be unambiguously classified as perimeter or lawnmower, and so were termed as 'other'.

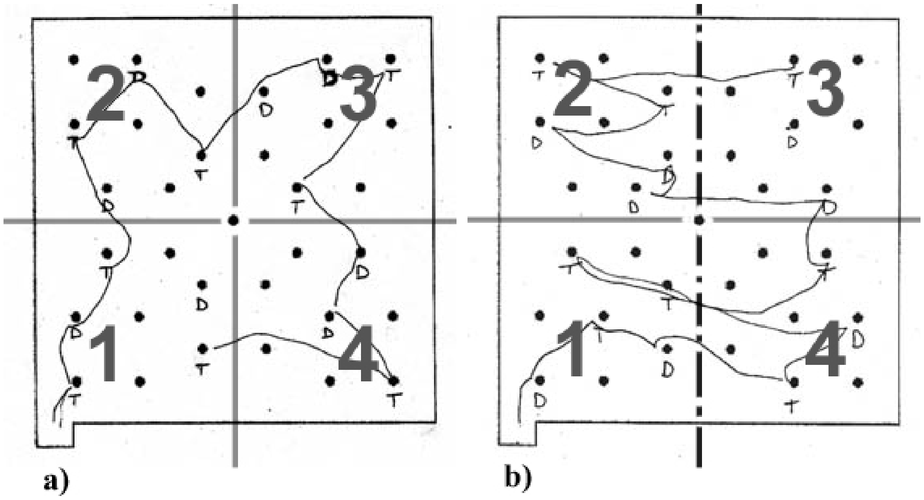


Figure 4.4: Two examples of participants' search paths in the real world experiment, a) perimeter search, and b) lawnmower search. Both figures show the plan divided into quadrants that were used for the classification of search strategies.

The second stage of the process was to count the number of times the centreline was crossed (the number of 'passes') during the primary phase. A perimeter search always had two passes, and a lawnmower search usually had three or more. The final stage involved counting the number of targets that were found by each search up until any target or decoy was revisited. The results for the three stages of the search strategy analysis are summarized in Table 4.1.

In the full-FOV condition, participants used a perimeter strategy for most of their searches, but a lawnmower strategy was dominant in the restricted-FOV condition. In all but one of the perimeter and lawnmower searches participants found all of the targets during the primary phase of the search. In the lawnmower searches participants made an average of four passes of the environment (up, down, up and then down again) compared with two for the perimeter searches. When using a lawnmower strategy participants tended to focus on a narrow "strip" of the environment during each pass, but with the perimeter strategy participants deviated

Table 4.1: Number of searches carried out with each strategy, mean number of targets found, and mean number of passes performed.

Group	**No. searches**			**Mean no. of targets found before repetition**			**Mean No. of passes**		
	Perim.	**Lawn.**	**Other**	**Perim.**	**Lawn.**	**Other**	**Perim.**	**Lawn.**	**Other**
full-FOV	11.00	3.00	1.00	7.91	8.00	8.00	2.00	3.70	2.00
restricted-FOV	4.00	10.00	1.00	8.00	7.70	5.00	2.00	4.20	2.00
both groups	15.00	13.00	2.00	7.93	7.77	6.50	2.00	4.10	2.00

from the edge of the environment to visit the targets and decoys that were nearby. The distance that participants travelled in excess of the shortest route was lower for trials performed using a perimeter strategy (M = 7.5%, SD = 15.8) than a lawnmower strategy (M = 19.2%, SD = 16.0).

4.4 Discussion

Restriction of the participants' FOV had no effect on participants' search performance in this real world cluttered space, but did affect the strategy they adopted. Two clear types of strategy were chosen by the participants (perimeter and lawnmower), with the perimeter strategy being dominant in the full-FOV condition, and the lawnmower strategy dominant in the restricted-FOV condition. One explanation could be that by reducing their FOV to such an extreme (20 x 16 degrees) these participants were forced to consider only nearby cylinders and were unable to plan an efficient route through the environment by considering the space as a whole. The resulting lawnmower strategy increased the number of changes in direction made by the participants throughout the trial, compared to the perimeter search, but because the real world offers such a rich source of proprioceptive and visual orientation cues, the restricted-FOV group did not become disorientated, and so did not visit significantly more targets and decoys than the full-FOV group. Overall and with both strategies, participants found the experimental task trivial to conduct, completing it with near perfect efficiency.

Two note-worthy comparisons may be made with the results of Ruddle and Jones (2001), who compared the performance of participants who used either a 48 or 103 degree FOV in a desktop VE to perform a task similar to the one used in the present real-world experiment. First, restricting a participant's real-world FOV had negligible effect on the time it took to complete the task, but in a desktop VE this restriction increased the time by approximately 40%. Second, restricting the real-world FOV had little effect on the number of targets and decoys visited, but in a VE caused a three-fold increase in the percentage of trials where participants had great difficulty completing the task and had to revisit at least half of the environment.

The experiment reported here investigated real world navigation by degrading real world sensory information (reducing the FOV). The majority of the full-view participants searched the environment with a perimeter strategy in a clockwise or anticlockwise direction, while the majority of the restricted-FOV participants adopted a lawnmower strategy. The restriction of participants' FOV did not increase the number of targets and decoys that were visited, compared to the full-view condition, indicating that participants were able to maintain their orientation throughout the task with the limited FOV and the use of vestibular and proprioceptive feedback.

Even though the lawnmower strategy created a longer search path with more turns of direction, it did not produce a decrease in performance. There were no improvements in performance across trials in either condition, indicating that participants were performing at ceiling level throughout the experiment. In both conditions, using either strategy, participants performed the search task with near perfect efficiency and found the task to be trivial.

All the participants who adopted the lawnmower strategy walked a route that was predominately in line with the structure of the cylinders and the walls of the sports hall: no one walked a lawnmower path that progressed across the cylinder layout at 45 degrees to the surrounding environment. This suggests that the participants were using the frame of reference of the cylinders and/or the sports hall as a guide (Mou & McNamara, 2002).

Chapter 5

Visual Fidelity and Movement Control

5.1 Introduction

As hypothesized, participants found the search task trivial to perform in a real world environment (see Chapter 4) whereas the same task has been found to be very difficult to perform in a VE. The experiment reported in this chapter was conducted to bridge the gap between the original VE study (Ruddle & Jones, 2001) and the real world environment by investigating the effect of different movement interfaces and visual scene characteristics on participants' search performance in a desktop VE.

The experiment used a 2 by 2 between participants design. Participants were randomly allocated to one of four conditions that each used one of two movement interfaces and one of two visual scenes. With one movement interface participants could only travel forwards (the *forward-only* condition), but the other allowed the participants to travel any combination of forwards, backwards, left and right (the *4-way* movement condition).

One visual scene condition used a VE with a high resolution "photorealistic" model of the sports hall (the *high-fidelity* scene condition). The other used the cylinder environment without a background (the *low-fidelity* scene condition) and was equivalent to the environments used in the original VE study. The high-fidelity scene condition was visually similar to the real world sports hall and used textures captured by digital camera from the floor, walls, and ceiling of the sports hall used in the real world experiment (see Figure 3.3). All four experimental conditions were implemented using desktop VEs.

5.2 Method

5.2.1 Participants

Twenty participants (9 females and 11 males) took part in the experiment, and their ages ranged from 18 to 40. Two female participants failed to finish the experiment. One participant failed to complete a single practice trial, even after an extended tuition period of 50 minutes, and so did not progress through to the test trials. Another participant withdrew due to symptoms of VE sickness. Presented here are the results of the 20 participants who successfully completed the experiment, including one participant who suffered from nausea in the last test trail but still finished the experiment. All participants were either graduates or undergraduates who volunteered for the experiment, were paid an honorarium for their participation, and gave informed content.

5.2.2 Materials

The experiment used a modified version of the software used in the original Ruddle and Jones (2001) study, adapted to reflect the changes made to the decoy and target boxes in the real world environment.

Participants travelled around the VE using the keyboard cursor key(s) for movement across the horizontal plane. Motion was continuous and at a speed of 1 metre/second while the cursor key(s) were depressed, and stopped when released. In the forward-only condition (view-direction movement) participants could only travel forward in the direction in which they were looking, achieved by holding down the ‘up’ cursor key. In the 4-way (independent movement) condition they could move forward, back, left, and right across the horizontal plane of the VE and could also travel diagonally by holding down pairs of keys (e.g. forward and left). In both of these movement interfaces the mouse was used to control the participant’s direction of view. Looking up and down was achieved using zero order control (i.e. position control, the movement of the mouse controlled the view directly); by moving the cursor up the screen the viewing pitch increased by up to +90 degrees, and the viewing pitch decreased by a corresponding amount if the cursor was moved down the screen. Looking left and right was accomplished using first order control (i.e. rate control, the movement of the mouse controlled the rate at which the view turned); the rate of turning increased proportionally with the cursor’s distance away from the vertical centreline of the screen. The maximum rate of turning was 135 degrees/second.

Participants could raise and lower the lids of the target and decoy boxes by pressing the left mouse button. Pressing it once raised the lid, and pressing it again lowered the lid. If there was a target present inside the box the participant pressed the right mouse button to select it, the target then turned from red to blue indicating that it had been found, comparable to the depositing of a blue card on top of a red target in the previous real world experiment. The lid was then lowered automatically by the VE software. The software prevented participants from moving away from any box until its lid was lowered. A lid could only be raised, and a target selected, if the participant was within 0.6 metres of the centre of the box (i.e., they were adjacent to it) and it was within the participant's FOV.

The VE application was written in C++ and OpenGL Performer and ran on a SGI Onyx 3400 with a constant frame rate of 60 Hz for all conditions, giving an overall system latency of approximately 30 ms. The VE was viewed with a 48 x 39 degree graphical FOV via a CRT colour monitor with a 475mm x 300mm viewable screen size; participants' physical FOV was therfore 43 degrees (horizontal) x 28 degrees (vertical). The resolution was 1280 x 1024 pixels and the refresh rate was 72Hz. All participants viewed the monocular displayed VE from a distance of approximately 60cm.

The high-fidelity scene condition used images of the sports hall surfaces as textures, these were captured using a digital camera and 'stitched' together to create seven separate textures: each of the 4 wall textures were 1024 x 512 pixels (1.5 Mb each in RGB format); the floor texture was 2048 x 1024 pixels (6.1 Mb), and the ceiling texture was 512 x 128 pixels (0.4 Mb). The seventh texture was used to replicate the appearance of the 25 lights, suspended from the sports hall ceiling (0.1 Mb each).

5.2.3 Procedure

The experiment followed the procedure outlined in Chapter 3. Participants were randomly allocated to one of the four conditions, were run individually, and took approximately 45 minutes to complete the experiment. Each participant first practiced using the interface, performed two practice trials, and then completed four test trials. Each trial began at the starting point in the boundary recess (see Figure 3.2), and participants were asked to minimize their journey path and to avoid checking each possible target location more than once.

5.3 Results

Participants' performance in each trial was measured using the same task performance metrics as Experiment 1, but additional behavioural metrics were used:

1. Task performance
    (a) Time taken to find the eight targets
    (b) The number of revisited target and decoy boxes, and the total number of visits to target and decoy boxes
    (c) Distance travelled
2. Behaviour
    (a) Movement key usage
    (b) Search strategy
    (c) Errors

Statistical analyses of the data followed the same method as the real world experiment and were performed using mixed design ANOVAs that treated the scene (high- vs. low-fidelity) and movement interface (forward-only vs. 4-way) as between participants factors, and the trial number as a repeated measure.

5.3.1 Task performance

Participants performed the searches significantly quicker with forward-only movement than with 4-way movement ($F(1,16) = 5.93$, $p < .05$), and significantly quicker in the high-fidelity scene than the low-fidelity scene ($F(1,16) = 8.16$, $p < .05$). Participants also performed the searches significantly quicker as the trials progressed ($F(3,48) = 2.93$, $p < .05$), with most of the difference between the conditions occurring in Trials 1 and 2 (see Figure 5.1).

Participants revisited fewer targets and decoys with the high-fidelity scene than the low-fidelity scene (M (SD) = 1.95 (3.52) vs. 4.4 (5.34); $F(1,16) = 6.13$, $p < .05$). Participants also revisited fewer boxes as the trials progressed ($F(3,48) = 2.84$, $p < .05$).

In terms of statistical differences, the pattern of results for the total number of visits to target and decoy boxes, and the percentage distance travelled above the minimum were identical to the time data. Participants visited fewer targets and

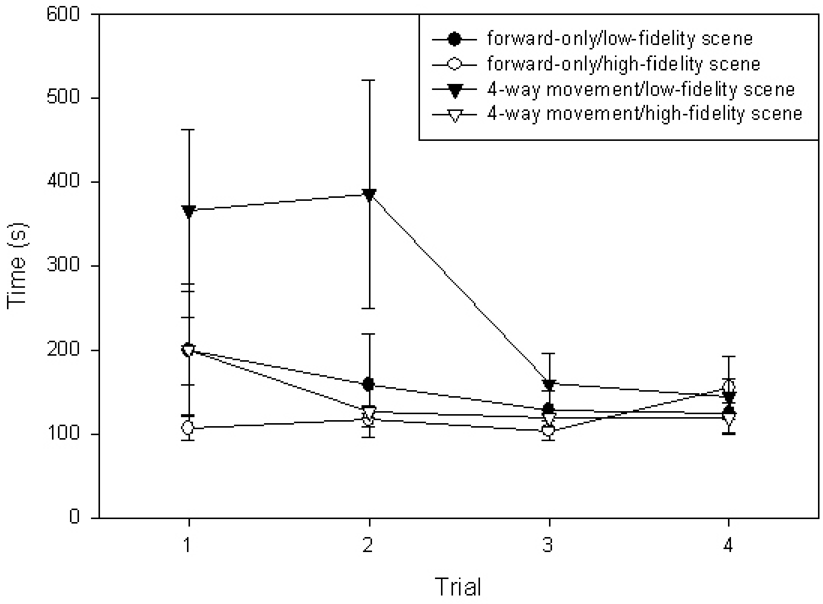


Figure 5.1: The mean times for the four VE movement/fidelity conditions. Error bars indicate the standard error.

decoys with the forward-only interface than the 4-way interface (M (SD) = 18.0 (6.4) vs. 22.6 (12.2); F(1,16) = 6.17, $p < .05$), fewer with the high-fidelity scene than the low-fidelity scene (M (SD) = 17.8 (6.6) vs. 22.8 (12.0); F(1,16) = 7.45, $p < .05$) and fewer as the trials progressed (F(3,48) = 3.10, $p < .05$). Participants traveled shorter distances with the forward-only interface than the 4-way interface (M (SD) = 44.2% (71.7) vs. 84.1% (104.0); F(1,16) = 4.55, $p < .05$), shorter distances with the high-fidelity scene than the low-fidelity scene (M (SD) = 41.5% (60.0) vs. 86.8% (110.1); F(1,16) = 5.87, $p < .05$) and shorter distances as the trials progressed (F(3,48) = 2.80, $p = .05$).

5.3.2 Behaviour

Three behavioural measures were used. First, the VE software automatically recorded the amount of time participants held down each of the keys that were used to control movement. Overall, participants held down a movement key for 12.8% of the trial time (excluding time when a box lid was raised and participants were prevented from changing position), and in this there was little difference between the four combinations of scene and movement interface, or the four trials. However, there was a marked difference between use of the 4-way movement interface with the two visual scenes. With the high-fidelity scene, participants used the forward key for 99.7% of the time they spent moving, and the left and right keys for the remaining 0.3%. In fact, three of the five participants in this group never used the left or right key. With the low-fidelity scene, participants used the forward key for 71.1% of the time, the backward key for 12.0%, and the left and right keys for 16.9%.

Second, participants' primary search strategies were classified using the same process as in the real world experiment (in fact, the process was developed by simultaneously looking at the data for both experiments). A perimeter strategy was dominant in the 4-way movement/low-fidelity group; the perimeter strategy was also dominant in the real world full-FOV condition (see Chapter 4). This is a surprising result because one could expect the 4-way movement/high-fidelity group to be dominated by the same (perimeter) strategy as the real world full-FOV condition since it allowed participants to move easily in many directions and looked most like the real sports hall. However, lawnmower and perimeter strategies were equally prevalent in the other three groups (see Table 5.1). The percentage distance above the optimum distance for the participants who used the lawnmower strategy (M = 40.0%, SD = 61.2) was almost half that of the participants that chose the perimeter strategy (M = 78.2%, SD = 107.7). This is in direct contrast to the findings in the real world experiment. Although participants in all four groups took a similar amount of time to perform the task in Trials 3 and 4, inspection of the number of targets missed during the primary search shows that no forward-only/high-fidelity participant ever missed more than one target, but at least one participant in each of the other groups missed three or more targets.

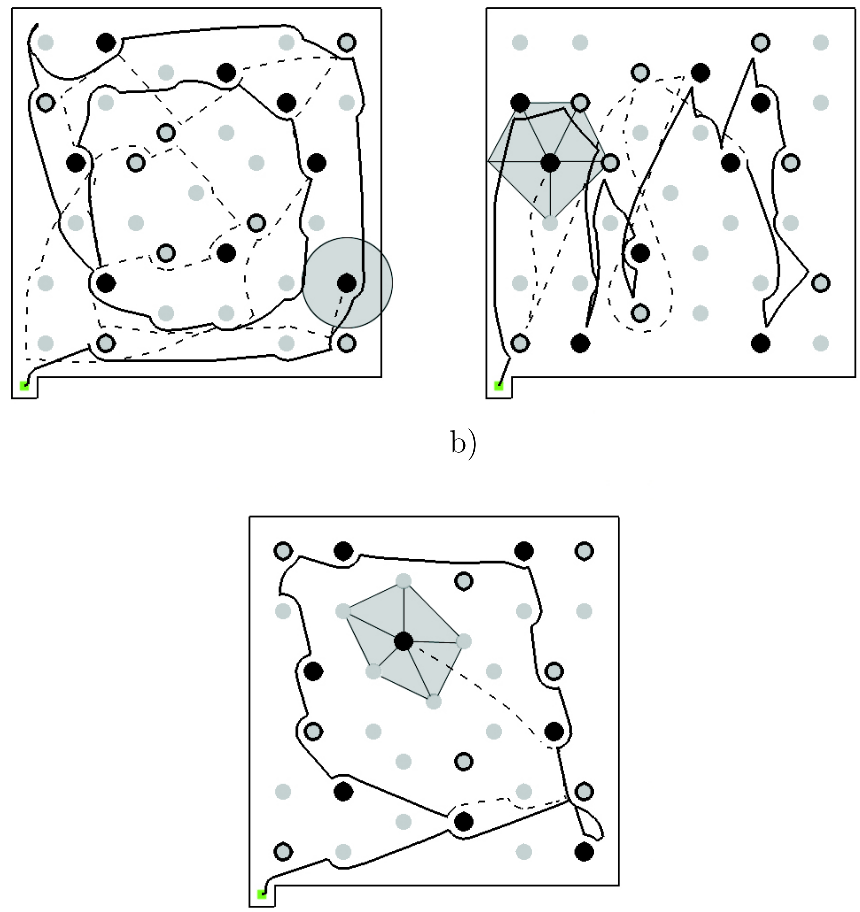
Table 5.1: Number of searches carried out with each strategy, mean number of targets found, and mean number of passes performed.

Group	**No. searches**			**Mean no. of targets found before repetition**			**Mean No. of passes**		
	Perim.	**Lawn.**	**Other**	**Perim.**	**Lawn.**	**Other**	**Perim.**	**Lawn.**	**Other**
forward-only/low-fidelity	8.0	9.0	3.0	7.13	7.33	5.67	2.0	3.6	3.3
forward-only/high-fidelity	9.0	10.0	1.0	7.44	8.00	7.00	2.0	3.9	2.0
4-way/low-fidelity	15	3.0	2.0	6.60	5.67	3.50	2.0	3.3	3.0
4-way/high-fidelity	11.0	8.0	1.0	6.09	7.88	3.00	2.0	4.4	2.0
All groups	43	30	7	6.74	7.53	4.86	2.0	3.9	2.8

In many of the trials, participants travelled substantially further than they needed to, revisiting many targets and decoys. Close inspection of the data showed that participants typically quickly found the first seven targets but then had difficulty finding the eighth. This is borne out by the fact that participants travelled an average of 6.1 metres to find each of the first seven targets but 18.5 metres for the eighth. The cause of these difficulties was errors made by participants, which forms the basis of the third behavioural measure. For each trial, we classified the targets that were found after one or more targets or decoys had been revisited into three groups by overlaying the path a participant had followed until the first revisit onto a plan view of the environment that had been divided into sectors using Delaunay triangulation (see Figure 5.2). A miss was recorded if the participant had previously touched the cylinder on which the target's box was located. Local neglect was recorded if the participant had previously travelled through any of the Delaunay triangles connected to the target's cylinder. Global neglect was recorded for all other errors, indicating that the participant had not been in the target's immediate vicinity. Overall, global neglect was prevalent in the forward-only/low-fidelity condition, but local and global neglect occurred with roughly equal frequency in the other conditions (see Figure 5.3).

5.3.3 Comparison with the real world experiment

A one-way ANOVA was performed to compare the mean number of visits that participants made to targets and decoys in the test trials of the two experiments. For the analysis, participants in both conditions of Experiment 1 (full-view and restricted-FOV) were combined into a single group because their performance had been almost identical. Overall there was a significant difference between the real-world and VE participants ($F(4,25) = 9.15$, $p < .01$). Planned contrasts showed that the real-world participants made significantly fewer visits than participants in the forward-only/low-fidelity, 4-way/low-fidelity, and 4-way /high-fidelity VE groups ($p < .05$), but not the forward-only/high-fidelity group. Detailed inspection of the data showed that, although the real-world and VE forward-only/high-fidelity groups visited a similar number of targets and decoys ($M = 15.3$ vs. 16.0), the latter only completed 55% of the trials perfectly (i.e. revisiting no targets or decoys) whereas the full-FOV real-world group were perfect on 93% of trials. With the exception of one of the imperfect trials, these VE participants never revisited more than two targets or decoys. The range of performance found in the four combinations of VE



a) b)

c)

Figure 5.2: Examples of the errors made by participants. The solid line shows the path up to the point that the first target or decoy was revisited, while the dashed line shows the participant's path for the remainder of the trial. a) Shows a missed target where the participant's path was deflected by the cylinder (see shaded circle). b) local neglect and c) global neglect show the shaded Delaunay triangulation area which defined the region through which the participant had to pass on their primary search if the error was to be defined as local neglect.

conditions is shown in Figure 5.4.

5.4 Discussion

The implementation of the movement interface had a significant effect on participants' search performance. Participants who used forward-only movement visited fewer targets, travelled a shorter distance, and took less time than participants who used the 4-way movement interface. These results echo the findings of Ruddle and

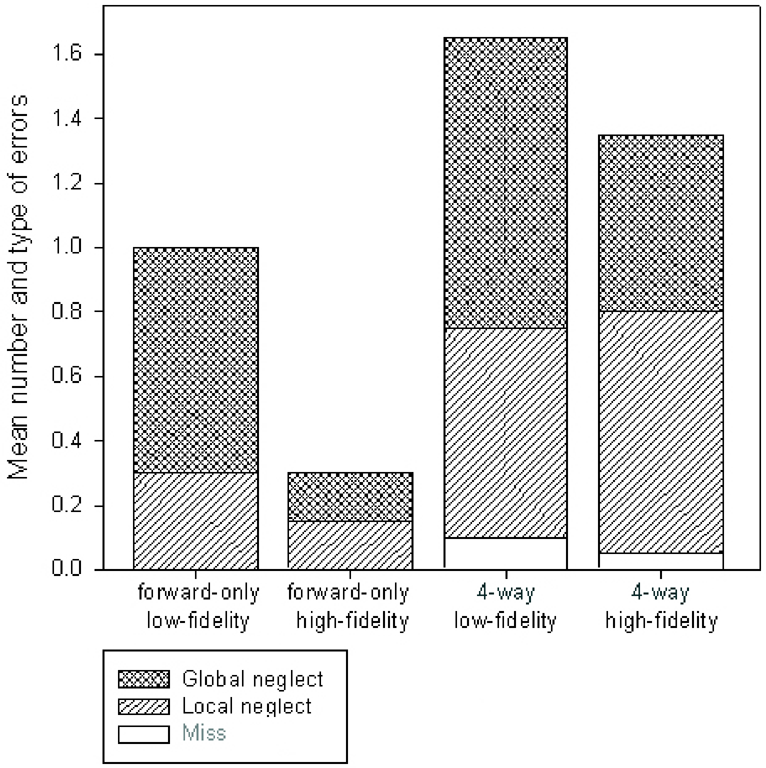


Figure 5.3: Mean number of each type of error made in each trial, for each combination of fidelity and movement interface.

Jones (2001) where the simplest movement interface was found to produce the most effective searches. However, by the third and fourth trial of the present study, participants achieved similar results with both forms of interface, but this improvement in performance was not due to a change in the type of primary search strategy that was adopted, as this rarely changed between trials.

The visual characteristics of the VE were also significant in affecting participants' ability to search the virtual space. The high-fidelity VE used large and detailed texture maps to create a visually faithful facsimile of the sports hall scene. This seems to have created a VE with adequate cues for the updating of orientation and heading across the test trials. It should also be noted that participants' familiarity with the real world sports hall varied, and was not controlled in the experiment. In the condition where the most effective movement interface was used in conjunction with the most effective visual scene, forward-only/high-fidelity scene condition, participant's efficiency (number of visited boxes) was similar to the real world.

Classification of the errors that participants made provides information about why participants searched inefficiently in many trials. Misses were rare and, as in

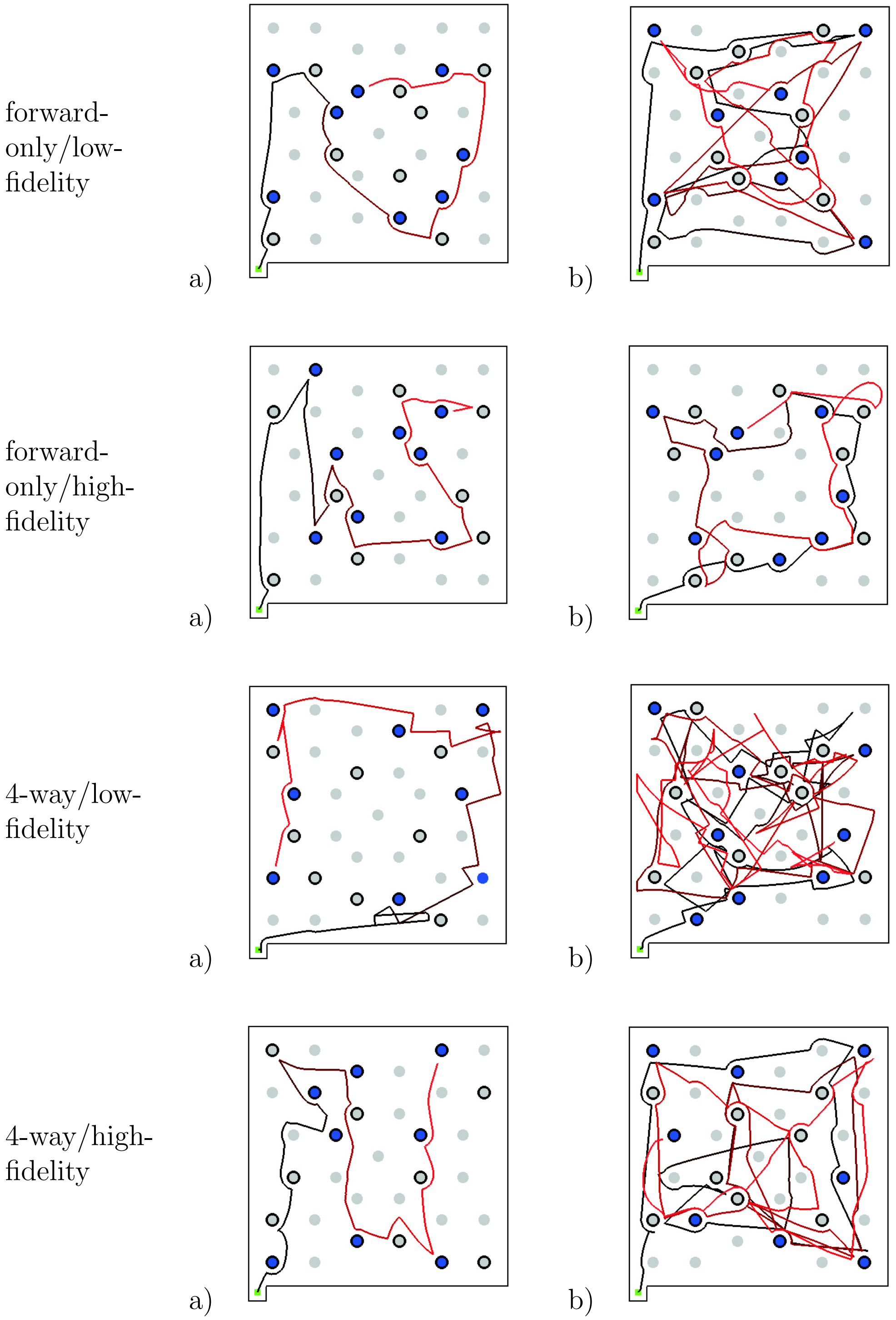


Figure 5.4: Best a) and worst b) search performance (revisits) for each VE movement and scene combination.

the study by Ruddle and Jones (2001), local and global neglect occurred with similar frequency. To prevent a miss, participants simply had to turn to face a given target and select it. To prevent local neglect, participants had to move a short distance across to the target, whereas prevention of global neglect involved participants in manoeuvring around the obstacles presented by other cylinders. Of particular interest is the fact that local neglect was most common with 4-way movement, despite the fact that this interface made it easiest for participants to move in any direction.

An unexpected finding was the different use of movement keys in the two 4-way conditions. Participants in the 4-way movement/high-fidelity condition used the back, left and right sideway keys for only 0.3% of the time spent travelling, while the participants in the 4-way movement/low-fidelity condition used these keys for 28.9% of the time. Both groups were shown and encouraged to use the keys in the same way, but participants in the high fidelity condition chose to perform the task by predominantly using the forward key. It is hypothesized that the lack of a dominant frame of reference for the low-fidelity scene led participants to rely on the movement keys to navigate around obstacles, thereby maintaining their global orientation. One participant took this to an extreme and adopted a novel movement method that only used the four movement keys for navigation and did not use the mouse at all. By contrast, with the high-fidelity scene, participants used the mouse to turn as they travelled forwards and used the scene content to maintain their orientation.

Chapter 6

Visual Fidelity and FOV

6.1 Introduction

The experiment reported in this chapter was conducted to help identify the thresholds of visual fidelity and FOV that are required for efficient VE navigation. Specifically, the visual scenes were a photorealistic model of the University of Leeds sports hall (high fidelity) and a model with identical geometry but which was rendered using a tiled brick texture (medium fidelity). The FOVs were normal for a VE (48 degrees; displayed on a single computer monitor) and wide (144 degrees; three monitors arranged in an arc that subtended the same angle). A 2 by 2 between participants design was used, with each participant searching the VE while using one combination of scene fidelity and FOV.

6.2 Method

6.2.1 Participants

Forty participants (20 females and 20 males) took part in the experiment and their ages ranged from 21 to 36. Two female participants suffered from simulator sickness, withdrew, and failed to finish the experiment. All the participants volunteered for the experiment and were paid an honorarium for their participation. Screening and monitoring of all participants were performed using the Short Symptoms Checklist (SSC)(see Appendix A), developed by the VIRART group (Cobb et al., 1999) (see Chapter 8 for sickness data), and all gave informed consent.

6.2.2 Materials

In all conditions participants travelled around the VE using a keyboard and mouse to interact. Participants used a view-direction interface that was the same as that used by the forward-only conditions in Experiment 2 (see Chapter 5), the geometry of the sports hall VEs was also the same.

The high-fidelity scene conditions used the same textures as the high fidelity conditions described in Chapter 5. However, although the floor texture was the same size, 2048 x 1024 pixels and 6.1 Mb (see Figure 3.5a), the sporting event lines were absent and so consisted of a repeated base texture. The medium fidelity VE used the same floor and ceiling textures but a repeated brick texture that had no other features for each of the four walls (see Figure 3.5b). Thus in the high fidelity VE all four walls of the sports hall were clearly distinguishable from each other (see Figure 6.1), but in the medium fidelity (brick tiles) VE all four walls looked similar (Figure 6.2).

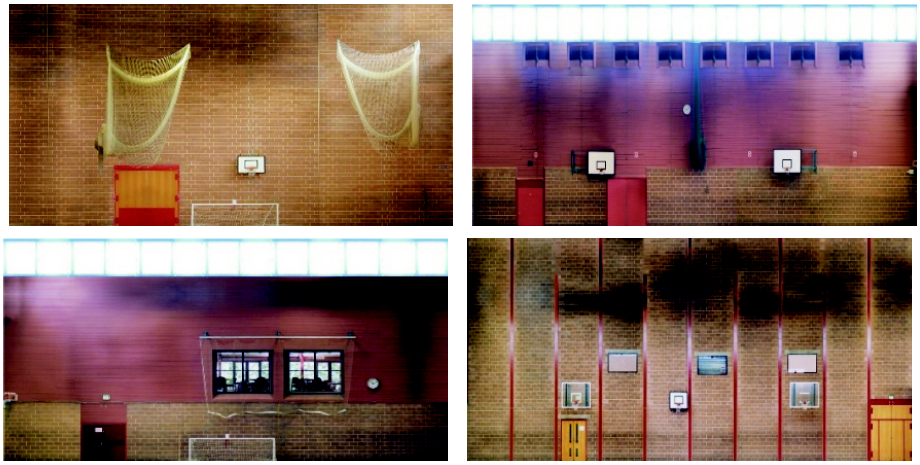


Figure 6.1: The four walls of the virtual sports hall (high fidelity scene).

In the normal-FOV condition participants viewed the VE via a monitor with a 475 mm x 300 mm viewable screen size. The resolution was 1280 x 1024 pixels and the GFOV was 48 degrees (horizontal) x 38 degrees (virtical). The physical FOV was 43 degrees (horizontal) x 28 degrees (vertical). The wide-FOV conditions used the same monitor as the normal-FOV conditions but with two additional monitors of the same type and size placed on its left and right. Each monitor displayed 1 of 3 viewing frustums at a resolution of 800 x 600 pixels. Taken together, the three monitors provided a continuous 144 degrees horizontal GFOV of the VE, with a physical FOV of aproximately 110 degrees (Figure 6.3). The difference in the views

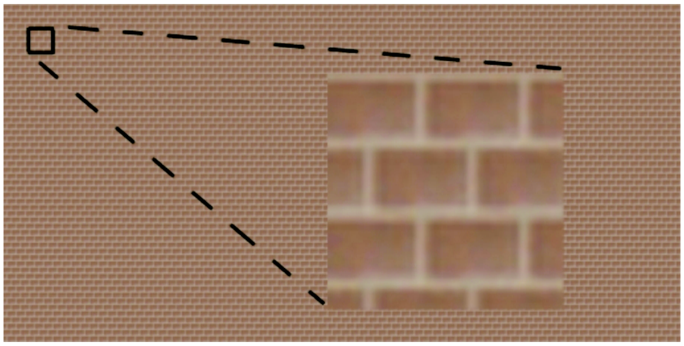


Figure 6.2: Close-up of the brick texture that was repeated (tiled) across all four walls of the medium fidelity VE.

seen by the participants with the two FOVs is shown in Figure 6.4.

6.2.3 Procedure

Participants were randomly allocated to one of the four conditions, were run individually, and took approximately 45 minutes to complete the experiment. Each participant first practiced using the view-direction movement interface, then performed two practice trials, and then completed four test trials. Each trial began at the starting point in the boundary recess (Figure 3.2) and participants were asked to minimize their journey path and to avoid revisiting possible target locations.

6.3 Results

Participants' performance in each trial was measured using a similar set of metrics to the previous study (Chapter 5):

1. Task performance
    - (a) Time taken to find the eight targets
    - (b) How close participants came to conducting a 'perfect search' (i.e. only checking each target and decoy once)
2. Behaviour
    - (a) The proportion of time spent performing different types of action
    - (b) Search strategy
    - (c) Errors

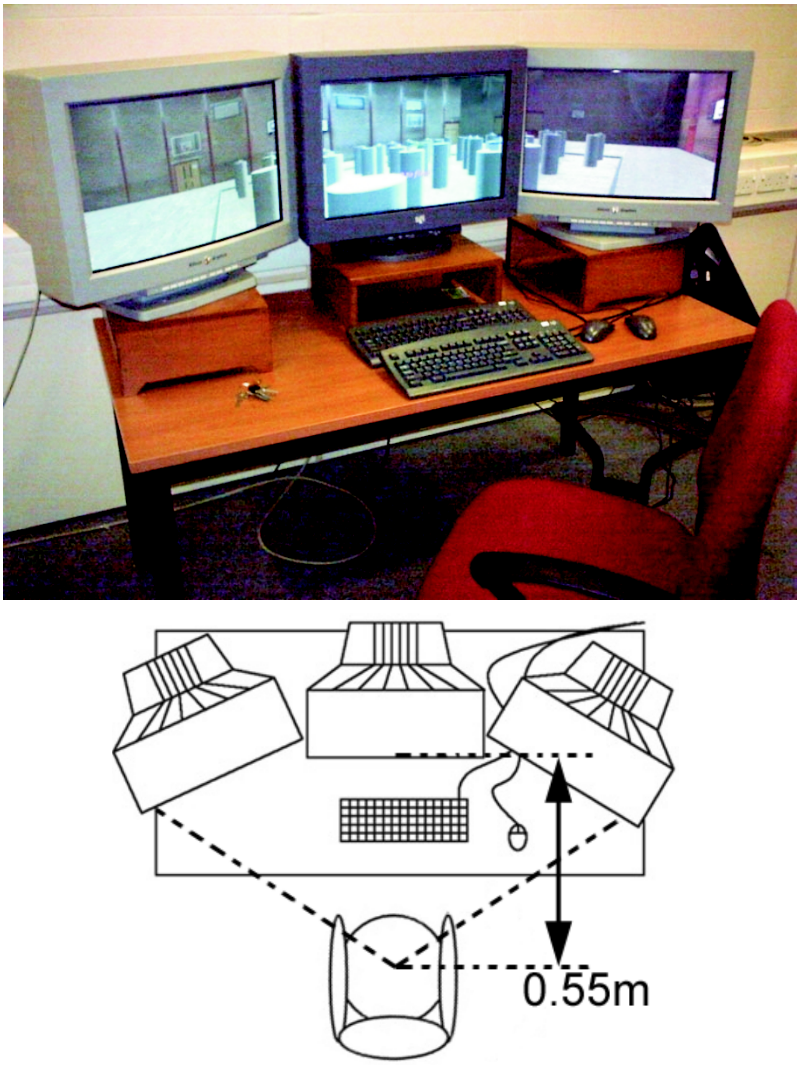


Figure 6.3: Photograph of the 3-monitor setup (above), and plan (below).

Statistical analyses of the data were performed using mixed design ANOVAs that treated the visual fidelity (brick tiles vs. sports hall) and FOV (normal-FOV vs. wide-FOV) as between participants factors, and the trial number as a repeated measure. Interactions are only reported if they were significant.

6.3.1 Task performance

Participants took less time to find the targets with the sports hall than with the brick tiled scene but the difference was not significant ($F(1, 36) = 2.21$, $p > .05$). Participants also took less time to find the targets with the wide FOV than with the normal FOV, but again the difference was not significant, ($F(1, 36) = 3.29$, p

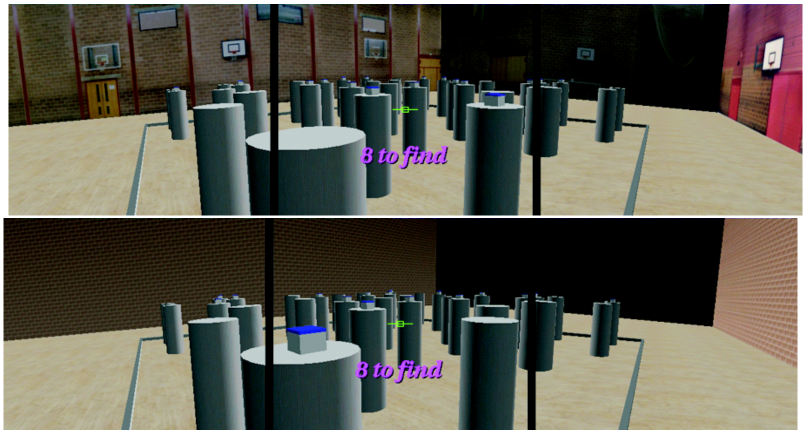


Figure 6.4: Views from each frustum for the wide FOV condition, high fidelity (above), medium fidelity (below). Only the middle frustum was visible in the normal FOV condition.

$> .05$). However, a significant learning effect occurred with participants taking less time to find the targets as the trials progressed ($F(1,\ 108) = 3.70$, $p < .05$), see Figure 6.5.

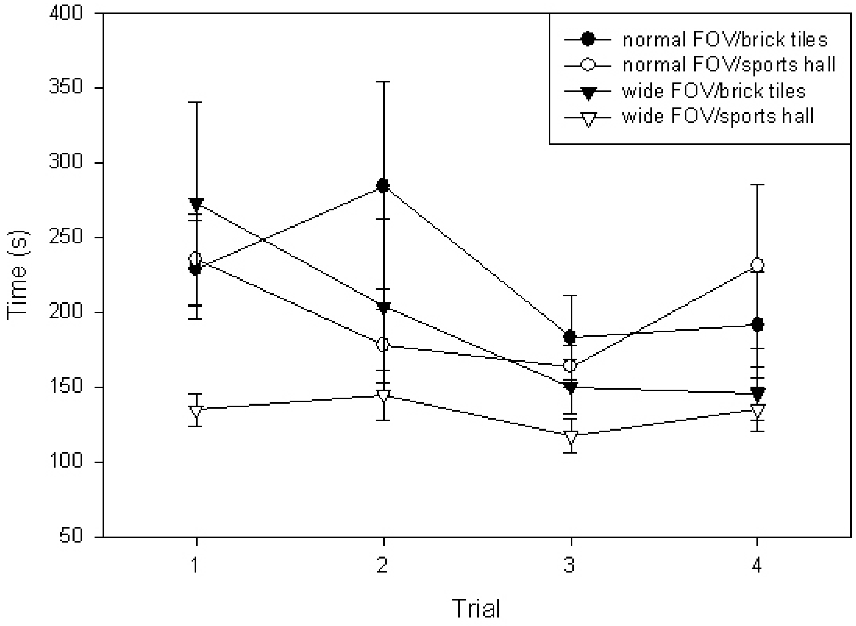


Figure 6.5: The mean times for the four FOV/fidelity conditions. Error bars indicate the standard error.

There were no significant differences for the number of revisited boxes between the normal FOV conditions and the wide FOV conditions (M (SD) = 3.28 (4.60) vs. 3.11 (4.39); $F(1{,}36) = 0.03$, $p > .05$). Also, no significant differences between the brick tiled conditions and the sports hall conditions (M (SD) = 3.66 (4.77) vs. 2.87 (4.22); $F(1,36) = 1.07$, $p > .05$).

When performing the task in the real world, participants completed 93% of the trials perfectly, visiting each target and decoy only once (see Chapter 4). However, in the present VE experiment only 51% of the trials were completed perfectly, and it was noticeable that participants often only found six or seven of the targets before checking again at least one target or decoy that had already been visited. A repeated measures ANOVA of the number of targets found before any revisitation showed a significant interaction between FOV and scene fidelity ($F(1, 36) = 5.24$, $p < .05$), with participants finding most targets with the wide-FOV/sports hall and least with the wide-FOV/brick tiles (M = 7.2 vs. 6.0).

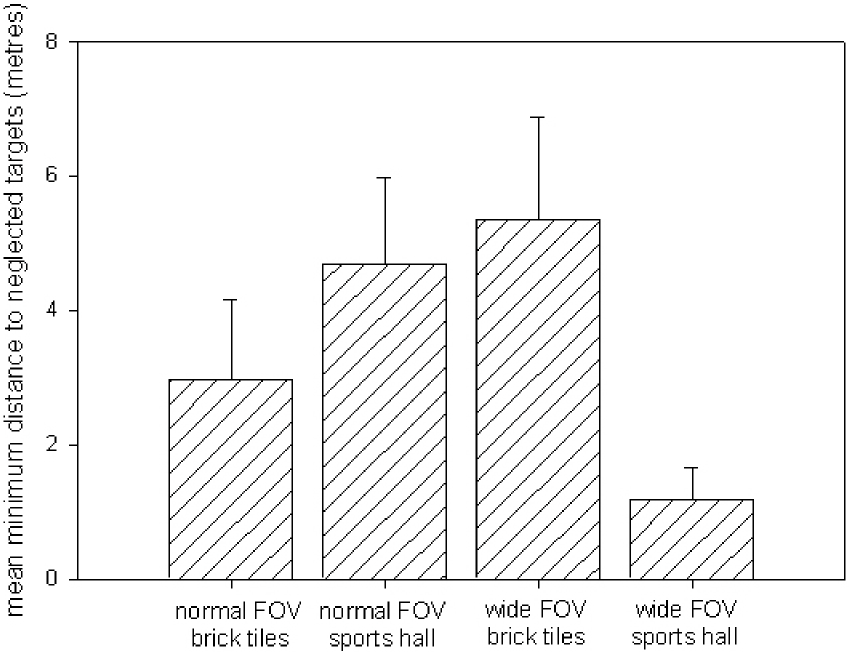


Figure 6.6: Sum of the minimum distance away from all unselected targets before the first revisitation occurred in a trial for the four FOV/fidelity conditions. Error bars indicate the standard error.

The number of targets found before any revisitation is a crude measure of how close a participant came to completing a trial perfectly, so a more sophisticated metric was developed that incorporated the proximity of a given participant to each target up until the first revisitation. Targets that had been found by this point scored zero, and the score for each remaining target was the closest distance (x) to the target so far (Figure 6.8). A repeated measures ANOVA showed a significant interaction between FOV and scene fidelity for the perfect search ($F(1, 36) = 4.40$, $p < .05$). Participants in the wide-FOV/sports hall condition came closest to a perfect search, whereas the wide-FOV/brick tiles group were the furthest away. This was a surprising result since one could expect the normal-FOV/brick tiles group to be further away from conducting a perfect search than the wide-FOV/brick tiles group, since they had a more restricted view of the VE, however, this was not the case, see Figure 6.6.

6.3.2 Behaviour

The first behavioural measure was the amount of time participants spent performing different types of action.

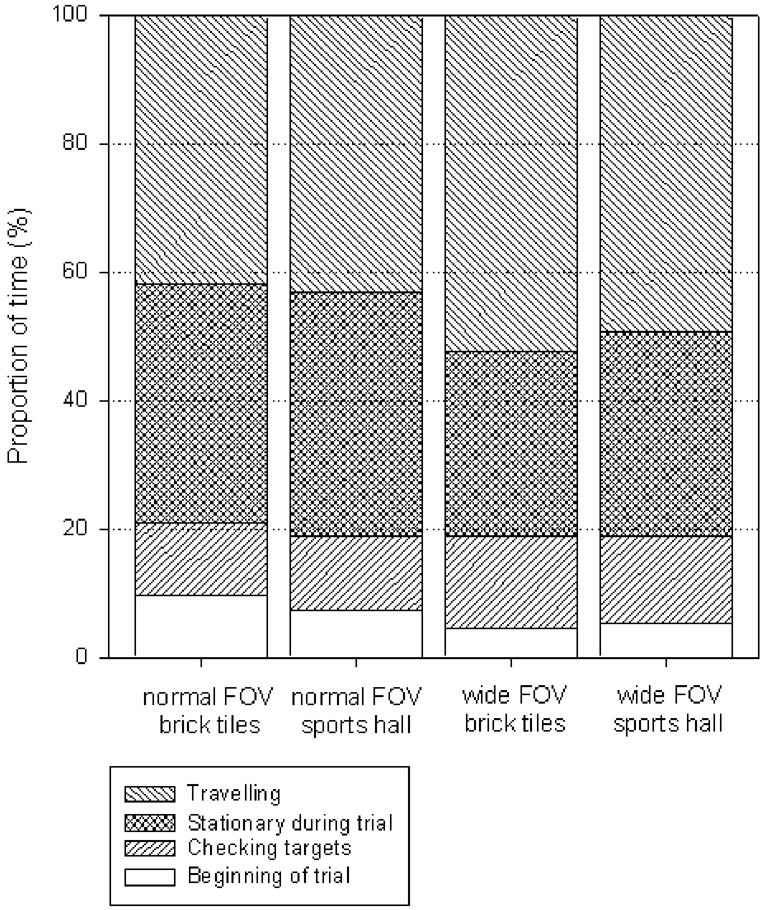
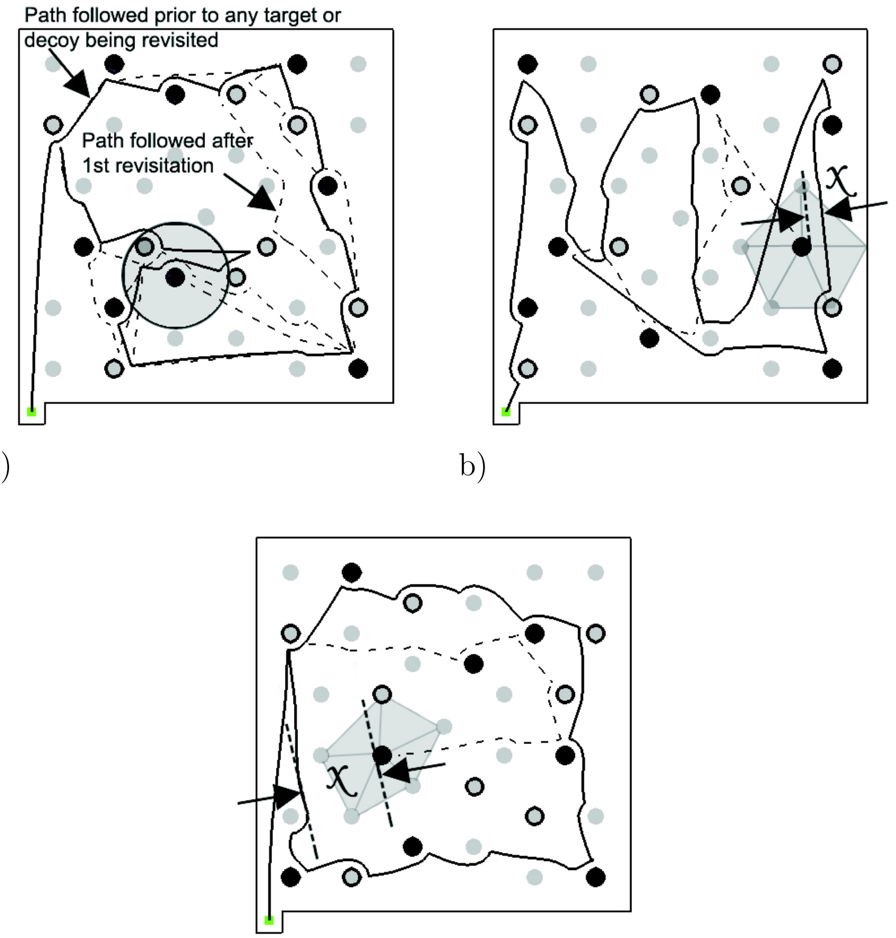


Figure 6.7: Proportion of time spent on various actions during a trial.

For each trial data were calculated for the percentage of time participants spent (i) planning where to go at the start of a trial, (ii) travelling around the VE, (iii) stationary between targets and decoys, and (iv) checking the target and decoy boxes. These data were then analyzed using repeated measures ANOVA's. Participants who used a wide FOV spent significantly less of their time planning ($F(1,\ 36) = 8.10$, $p < .05$), and stationary ($F(1,\ 36) = 4.47$, $p < .05$), than participants who used a normal FOV. It follows that participants who used a wide FOV spent significantly more of their time checking boxes ($F(1,\ 36) = 8.71$, $p < .05$), and travelling around than participants who used a normal FOV ($F(1,\ 36) = 5.88$, $p < .05$). (Figure 6.7).

The second behavioural measure was participants' overall search strategy. As in previous experiments the paths participants followed were classified by dividing the VE into four quadrants and noting the order in which these were visited (see Chapter 4, Figure 4.4).



a)

b)

c)

Figure 6.8: Examples of the three types of error a) miss, b) local neglect, and c) global neglect. 'x' shows the distance used in the 'perfect search' performance measure (x = 0 for a miss).

Two searches in the wide-FOV/brick tiled condition could not be unambiguously classified as perimeter or lawnmower, and so were termed as ‘other’. Approximately half of the participants in every group completed the task by using the perimeter strategy, while the other half used the lawnmower strategy. These are the same strategies as were used by participants who performed the same task in previous experiments.

To explain the difficulties that participants encountered, targets that were found after one or more targets or decoys had been revisited were classified into one of three categories, *miss* (Figure 6.8a), *local* (Figure 6.8b) or *global neglect* (Figure 6.8c), depending on a participant’s travelled path and its relationship to the neglected target (see Chapter 5 for a detailed description).

There was one miss in each of the normal FOV conditions, six in the wide-FOV/brick tiles, and four in the wide-FOV/sports hall. The ratio of local:global neglect was similar in all four conditions but, in total, in the wide-FOV/sports hall condition, participants made approximately half the number of errors than participants in both the normal-FOV/sports hall and wide-FOV/brick tiles conditions, and a third less than the normal-FOV/brick tiles condition (Figure 6.9). Overall, 5% of the targets were missed, 27% were locally neglected, and 68% were globally neglected.

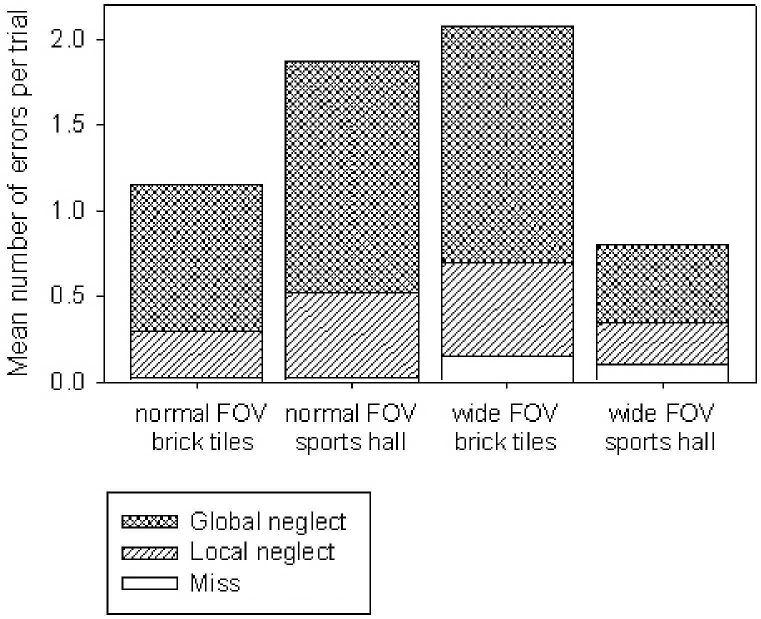


Figure 6.9: Type and mean number of errors for the four experimental conditions.

Further analysis of the path and direction of view of participants during trials revealed that, out of the 31 locally neglected targets in the normal FOV conditions, none actually came within the participants' FOV when they were in the targets immediate vicinity. This gives an explanation as to why so many targets fell into this category of error. It wasn't that participants neglected these targets, they simply didn't see them (see Figure 6.10). However, out of the 32 locally neglected targets in the wide FOV conditions, 17 were within the participants FOV, and all appeared in either the left or the right screen. Despite being visible and within close proximity to the participants, the targets remained unsearched.

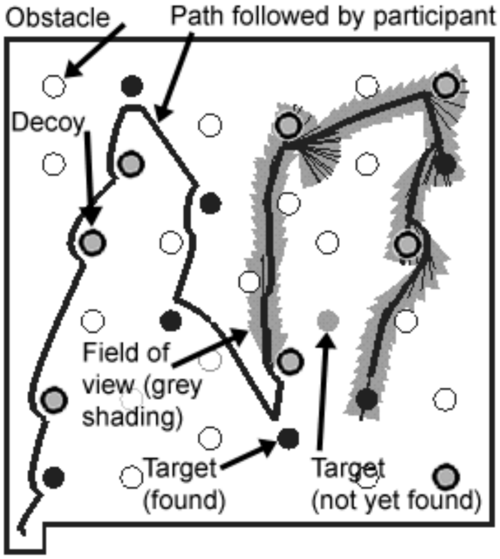


Figure 6.10: Plan view showing a participant's path as they travelled near to a target. Despite the proximity of the participant to the target it did not come within the participants FOV and so remained unsearched.

6.4 Discussion

This study investigated the effect of visual scene fidelity and FOV on participants' navigational ability in terms of task performance and behavioural metrics. With the combination of a wide FOV and a high fidelity (sports hall) scene participants were significantly closer to conducting "perfect searches" which is what is known to happen on the vast majority of occasions when people perform the same task in the real world.

Results from the behavioural metrics show significant differences occurred between the two FOVs. With a wide FOV participants spent proportionally less time standing in one place at the beginning of a trial and between target and decoys, than participants with a normal FOV. With a normal FOV, participants often stood in one place to look around and plan which box to search next, but a wide FOV allowed participants to assimilate the scene more easily simply because more of the VE was visible at any one time. It is reasonable to suggest that if more of the scene is outside the FOV then it is more demanding to keep track of the locations of the possible targets and this then requires more looking around.

It should also be noted that looking around while travelling with a view-direction movement metaphor causes participants to veer off course. This metaphor is the one that is most commonly implemented in VEs, and was the one used in the present study. By contrast, observations from the real-world study in which participants walked around a physical version of the environment (see Chapter 4) showed that participants rarely stopped between targets and occasionally did not even stop to check boxes.

Participants searched the VE using two strategies, perimeter, or lawnmower. Approximately half of the participants in every group completed the task by using the perimeter strategy, while the other half used the lawnmower. The change in FOV did not change the frequency of one strategy over the other, unlike the real-world study reported in Chapter 4 where the majority of participants who had a normal view of the environment searched the space with a perimeter strategy, while with a restricted FOV (20 degrees) the lawnmower strategy dominated.

The errors that prevented participants from conducting perfect searches were classified into three types, miss, local and global neglect. Participants in the wide-FOV/sports hall condition had far fewer instances of local and global neglect than the other three conditions.

Errors occurred for three main reasons. First, both local and global neglect could occur when participants only searched some of the targets and decoys that existed in a tight cluster, whereas in the real-world participants tended to exhaustively search clusters when they were encountered. Second, entire clusters of targets were sometimes globally neglected in all the VE conditions, indicating that even with a wide FOV and a high fidelity scene, participants often had difficulty remembering where they had travelled. Lastly, in the two normal FOV conditions all of the locally neglected targets were outside the participant's FOV when they were within the targets immediate vicinity. Therefore it is unlikely that participants were even aware of the targets locations despite being in close proximity to them. In the wide FOV condition 53% of the locally neglected targets were within the participants

FOV, and yet they were still not searched. Why were these targets neglected? It is possible that they were either not noticed or participants believed they had already been searched. If the former is true then it may have been caused by a difficulty participants had attending to the large physical area of the three-screen display. If the latter is true then one explanation might be that the participants where unable to assimilate the spatial information presented by the three separate screens into one useable representation of the sports hall. Even though more information was available they could not keep the positions of the targets updated with their own position and orientation as they travelled around the VE.

A surprising result was that participants in the normal-FOV/brick tiles condition came closer to conducting a perfect search, and made fewer errors, than those in the wide-FOV/brick tiles condition. The wider FOV seems to have hindered participants' ability to search the VE efficiently, although the reason for this remains unknown and requires further investigation.

In conclusion, the present study shows that a shift towards real-world navigational performance occurs when a wide FOV is combined with a photorealistic scene. However, even with this combination of factors, performance remains substantially below that which occurs in the real world. In turn, this implicates the likely importance of more flexible forms of movement for efficient VE navigation.

Chapter 7

Visual Fidelity and Physical Movement

7.1 Introduction

The experiments reported in this chapter investigated the effects of different movement interfaces and scene fidelity on participants' search performance. Since physical movement has been shown to improve spatial tasks (see Chapter 1) it was hypothesised that the more natural the interface, i.e. the more similar it is to natural walking, the better participants' performance. Therefore, physically walking around a VE while viewing that VE via a HMD (a "roving" HMD; physical translational and rotational movement) would produce better search performance results than a condition that required participants to remain stationary but allowed them to turn around to view the VE via the HMD (a "tethered" HMD; physical rotational movement but abstract translational movement). Also, the tethered HMD condition would produce better performance results than an interface where participants remain physically stationary (a desktop monitor condition; abstract rotational and translational movement). In the main experiment three conditions used a high visual fidelity VE, while in a supplementary experiment a roving HMD and a desktop condition used a low fidelity visual scene. The supplementary experiment was performed to investigate whether differences found in the main experiment persisted when visual scene fidelity was degraded.

7.2 Experiment

Participants were randomly allocated to one of the three high-fidelity visual scene conditions: desktop, tethered HMD, or roving HMD, or one of the two low fidelity scene conditions: desktop, or roving HMD. Participants in the desktop conditions searched the virtual space while seated in a chair and navigated with a keyboard and mouse. In the tethered HMD condition participants physically stood in one place and turned around to view the VE and travelled with the use of a handheld 3D mouse. Participants in the roving HMD conditions physically walked through real space while a tracking system updated their position and orientation in the VE, (Usoh, Arthur, Whitton, Bastos, Steed, Slater & Brooks, 1999).

In the high visual fidelity HMD conditions participants conducted the search task while present in the real world VR laboratory space that the VE depicted, the participants in the high visual fidelity desktop condition navigated the same VE but in a different real world space. The low visual fidelity roving HMD condition was conducted in the same VR laboratory but participants navigated around a low visual fidelity VE. Participants in the low visual fidelity desktop condition navigated the same low visual fidelity VE but in the same real world space as the high visual fidelity desktop condition.

7.3 Method

7.3.1 Participants

Fifty participants (24 males and 26 females) took part in the experiment, and their ages ranged from 18 to 32 (M = 22.9, *SD* 3.7). All participants were either graduates, undergraduates, or staff of the University of Leeds who volunteered for the experiment and were paid an honorarium for their participation and gave informed consent. Participants were screened before the experiment and monitored throughout all trials for signs of simulator sickness, six participants in the HMD conditions suffered mild symptoms (see Chapter 8).

7.3.2 Materials

This experiment used the same software as the previous experiments, but the dimensions of the boxes and the spacing and diameter of the cylinders were reduced to fit within the physical VR laboratory space. The position of the high visual fidelity

VE (see Figure 3.7a and b) matched the position of the real laboratory. The desktop conditions used the same movement controls as reported for the forward-only condition (view-direction movement) in Chapter 5. The HMD (a Virtual Research V8; see Figure 1.9) and tracking system (a WorldViz PPT and Intersense Inertiacube2) used for the tethered and roving HMD conditions are described in Chapter 1. Nothing in the roving HMD condition prevented participants from walking though the VE cylinders, so to make movement comparable in all five conditions collision detection was not performed with the cylinders in the desktop and tethered conditions. However, in the desktop and tethered conditions collision detection was used to prevent participants from moving outside the environment as a whole, just as the physical walls of the laboratory prevented this movement from occurring with the roving HMD. In both the desktop and tethered HMD conditions participants were prevented from moving away from a box until its lid was lowered, and in the roving HMD conditions participants were prevented from raising a lid until all were in the lowered position. In all conditions, a lid could only be raised, and a target selected, if the box was on the nearest cylinder to the participant and it was within the participant's FOV.

7.3.3 Procedure

Participants were run individually, and took approximately one hour to complete the experiment procedures in the desktop conditions, and two hours in the HMD conditions. The procedures for the conditions are outlined in table 3.1. All participants first practiced using the interface on the desktop until they could fluently use the controls. The controls were identical to the forward-only conditions (view-direction movement interface), see chapter 5. Participants then performed two practice trials in the virtual laboratory that allowed him or her to become familiar with the search task. Participants in the desktop conditions then completed two more practice trials and then completed four test trials. Participants in the three HMD conditions were taken into the real laboratory where they were shown the HMD and given a demonstration of the controls. Participants held a 3D mouse (see Figure 1.5), pressed one button to raise and lower a lid, and another to choose a target - the lid then lowered automatically. In the tethered HMD condition a third button was used for forward movement, participants could only travel forward in the direction in which they were looking. Once comfortable with the HMD the participants then practiced the controls, performed a further two practice trials, and then completed four test trails, using the tethered HMD or walking interface according to the condition to which a participant had been allocated.

Sickness data were recorded using a paper version of the Short Symptoms Checklist (SSC) (Cobb et al., 1999), see Chapter 8 and Appendix A for details. After each trial in the HMD conditions participants were asked to remove the HMD and sit down for a minimum of three minutes as a precaution against VE sickness. Before the next trial began participants were asked if they could feel any effects such as nausea or dizziness etc, the experiment continued only if they reported that they could not feel any symptoms. At the end of the four test trials participants were asked to remain in the department for one hour as a precautionary measure against the delayed onset of symptoms. During this hour participants indicated on the SSC the severity of any symptoms at 15 minute intervals, until the one hour had elapsed. Participants were asked to remain in the department if they suffered from any symptoms, until all symptoms had disappeared. Two participants in the tethered HMD condition and four in the roving HMD condition reported suffering from minor symptoms (see Chapter 8).

7.4 Results

Participants' performance in each trial was measured using a similar set of metrics as the previous experiments.

1. Task performance
    - (a) Time taken to find the eight targets
    - (b) Total number of revisited boxes
    - (c) How close participants came to conducting a 'perfect search' (i.e. only checking each target and decoy once)
2. Behaviour
    - (a) Search strategy
    - (b) Errors

Statistical analyses of the data followed a similar method as previous experiments and used mixed design ANOVAs that treated the movement interface (main experiment: desktop vs. tethered HMD vs. roving HMD; supplementary experiment: desktop vs. roving HMD) as between participants factors, and the trial number as a repeated measure. Interactions are only reported if they were significant.

7.4.1 Main Experiment - Performance

There were no significant differences in the time participants took to complete the trials between the three movement interfaces ($F(2,27) = 0.90$, $p > .05$), or as the test trials progressed ($F(3,81) = 1.18$, $p > .05$). It is noteworthy that participants' speed in the desktop and tethered conditions was a constant 0.5 metres per second, while participants could move as fast, or as sow, as they liked in the roving HMD condition.

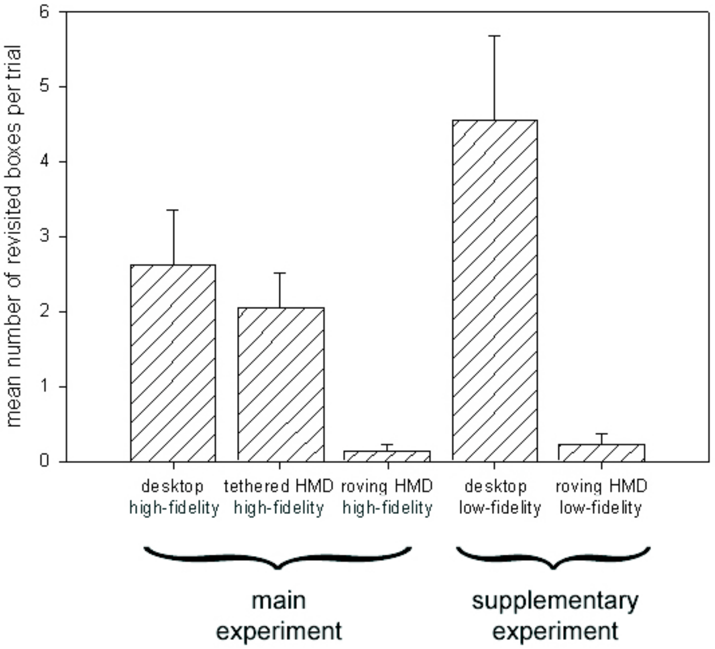


Figure 7.1: The mean number of revisited target and decoy boxes during test trials. Error bars indicate the standard error.

There were highly significant differences in the number of revisited target and decoy boxes during test trials ($F(2,27) = 6.60$, $p < .005$), but not between test trials ($F(3,81) = 0.47$, $p > .05$). Bonferroni-corrected post-hoc multiple comparisons showed that participants revisited significantly fewer target and decoy boxes in the roving HMD condition than the tethered and desktop conditions ($p < .05$). The difference between the tethered and desktop conditions was not significant (see Figure 7.1).

There were also highly significant differences between the number of targets found before a target or decoy box was revisited during test trials ($F(2,27) = 7.09$, $p < .005$), but not between test trials ($F(3,81) = 1.74$, $p > .05$). Bonferroni-corrected post-hoc multiple comparisons showed that participants found significantly more targets before revisiting a target or decoy box in the roving HMD condition than the tethered ($p < .005$), and desktop ($p < .05$) conditions. The difference between the tethered and desktop conditions was not significant (see Figure 7.2). The percentage of trials completed *perfectly* for the roving HMD, tethered HMD, and desktop conditions were 90%, 45%, and 42.5% respectively.

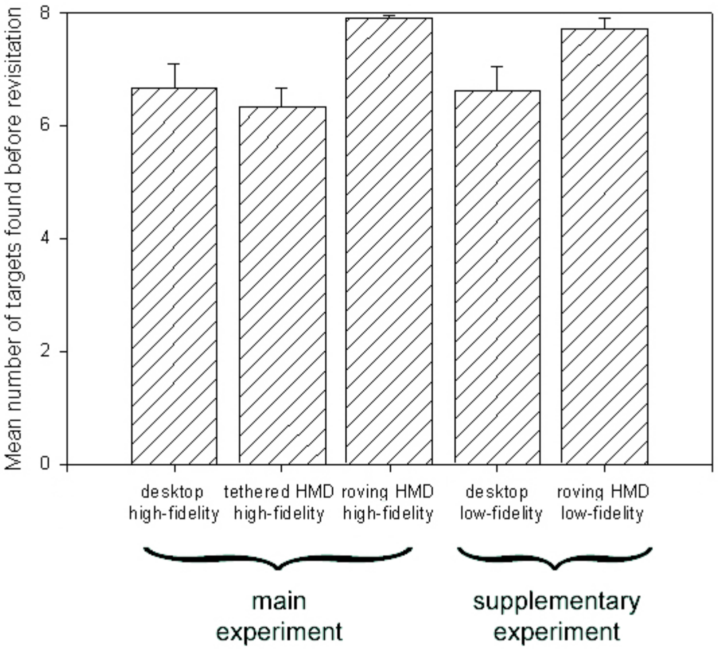


Figure 7.2: The mean number of targets found before a target or decoy box was revisited. Error bars indicate the standard error.

7.4.2 Main Experiment - Behaviour

There were more than twice the number of perimeter searches than lawnmower in the desktop condition (29 vs. 11), a similar result in the tethered condition (28 vs. 12), while a perimeter search dominated in the roving HMD condition (38 vs. 2). As in previous experiments errors that prevented participants from conducting a perfect search were classified into miss, local and global neglect. The number and type of errors in the three interface conditions are shown in Figure 7.3. In those trials where a participant was in the immediate vicinity of a neglected target (i.e. missed and locally neglected targets) a visual inspection of participants' travelled path was conducted to establish whether or not it had appeared inside the participants FOV (see Table 7.1).

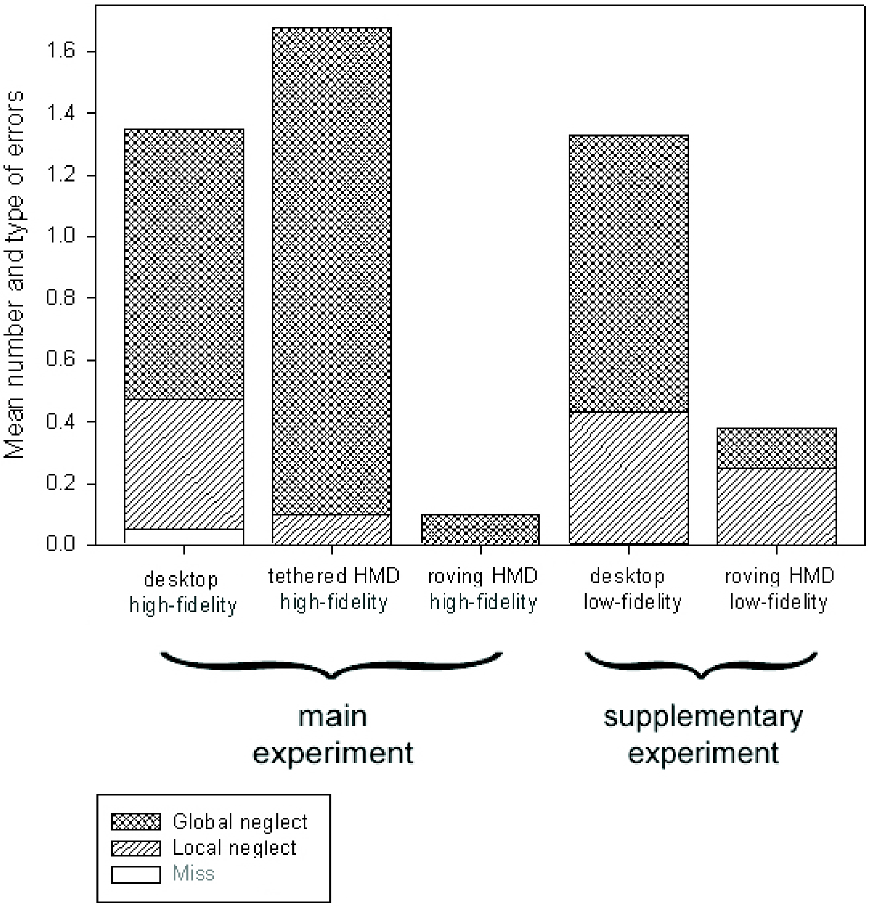


Figure 7.3: The mean number and type of neglected targets in the three movement interface conditions.

Table 7.1: Missed and locally neglected targets in or outside a participant's FOV.

		miss		local	
		inside FOV	outside FOV	inside FOV	outside FOV
Main Expt.	Roving HMD/high-fidelity	0	0	0	0
	Tethered HMD/high-fidelity	0	0	2	2
	Desktop/high-fidelity	2	0	14	3
Supp. Expt.	Roving HMD/low-fidelity	0	0	1	0
	Desktop/low-fidelity	2	0	10	7

7.4.3 Supplementary Experiment - Performance

There were no significant differences in the time participants took to complete the trials between the desktop and roving interfaces (M (SD) = 112.1 (72.47) vs. 83.57

(20.84); $F(1,18) = 1.68$, $p > .05$), or as the test trials progressed ($F(3,54) = 1.78$, $p > .05$).

There were highly significant differences in the number of revisited target and decoy boxes during test trials ($F(1,18) = 14.32$, $p < .005$) but not between test trials ($F(3,54) = 0.76$, $p > .05$) (see Figure 7.1).

There were also significant differences between the number of targets found before a target or decoy box was revisited during test trials ($F(1,18) = 5.82$, $p < .05$), but not between test trials ($F(3,54) = 1.46$, $p > .05$), (see Figure 7.2). The percentage of trials completed *perfectly* for the roving HMD and desktop conditions were 90% and 45%, respectively.

7.4.4 Supplementary Experiment - Behaviour

In the desktop condition there were more than twice the number of perimeter searches than lawnmower (26 vs. 12), with two that were classified as *other*, while the number of perimeter and lawnmower searches in the roving HMD condition were similar (23 vs. 17). As before, the errors that prevented participants from conducting perfect searches were classified into miss, local and global neglect. The number and type of neglected target in the two interface conditions are shown in Figure 7.3. A visual inspection was used to establish whether or not a target had appeared inside the participant's FOV (see Table 7.1).

7.5 Discussion

The main experiment used a high fidelity visual scene and showed the method of movement had a dramatic, and highly significant, effect on participants' performance. Participants that physically walked around the VE performed the task as efficiently as participants who took part in the real word experiment described in Chapter 4. Participants who walked revisited fewer boxes, and found more targets before revisiting boxes, than participants in both the tethered and desktop conditions. Surprisingly, there were no significant differences between the tethered and desktop conditions. Research investigating simple tasks such as direction estimates (Presson & Montello, 1994; Rieser, 1989) and angles turned (Bakker et al. 1999) suggest that there are benefits from physical rotation (tethered HMD) over visual turning (desktop) while physically stationary. However, in the present study, for this somewhat more complex spatial task, the tethered HMD interface gave participants no performance benefits over the stationary desktop display.

The percentage of trials conducted perfectly also reveals the fundamental difference in performance between the walking and the two other conditions. With the roving HMD 90% of trials were conducted perfectly, without researching any boxes, while in the tethered and desktop conditions only 45%, and 42.5% of trials, respectively, were conducted perfectly. Classification of the search strategies used by participants in the three high fidelity conditions revealed that in the desktop and tethered conditions participants used a perimeter strategy in only a third of the trials, whereas in the roving HMD condition a perimeter search strategy was used in all but two of the trials. However, the strategy chosen did not cause the difference between the number of searches that were perfect.

The supplementary experiment contained two movement conditions, walking and desktop, and both used a low fidelity visual scene. A reduction in the fidelity of the visual scene had little impact when participants walked, extending the findings of studies of path integration to a more complex setting (Kearns et al., 2002). However, with the desktop display a reduction in visual fidelity did lead to a reduced performance, in line with findings reported earlier in this thesis.

In conclusion, participants that physically navigated around the space with a space-consistent display, the roving HMD condition, dramatically out-performed those that navigated while stationary using a space-constant display device i.e. the desktop monitor. The highly significant performance differences between the walking and the two other interfaces suggests that full proprioceptive information combined with vestibular feedback dramatically improves search task performance in a small-scale cluttered environment.

Chapter 8

VE Sickness Data

8.1 Introduction

Presented here are the data of participants self reported symptoms of VE sickness for the three VE studies. Taken as a whole, 90 participants took part in desktop VE experiments, five (5.6%) suffered from some form of VE sickness, three of whom withdrew. Out of the 30 participants that took part in the immersive HMD experiments six (20%) suffered from symptoms of VE sickness, the symptoms were minor, quickly disappeared, and no one withdrew.

8.2 Visual Fidelity and Movement Control

It was not expected that participants would suffer from VE sickness while navigating via the desktop display therefore participants were not screened or monitored, VE sickness details were therefore not recorded, however two participants did suffer symptoms. One participant complained of dizziness and nausea during the early stages of the four test trials and the experiment was immediately stopped. One other participant reported mild nausea after the fourth test trial, reported that it had not been severe enough to influence their performance, and was able to finish the experiment. It should be noted that participants navigated the VE whilst viewing a large monitor with a viewable screen size of 475 x 300 mm (approximately 23 inches measured diagonally), and this could have contributed to the unexpected cases of VE Sickness.

8.3 Visual Fidelity and FOV

Due to the two participants who reported symptoms of VE sickness in the above experiment, screening and monitoring of participants in all subsequent experiments were performed using the Short Symptoms Checklist (SSC) developed by the VIRART group (Cobb et al., 1999) (see Appendix A).

Participants were asked to initial a statement to confirm that they did not currently suffered from any of the following: hay fever, migraines, heart conditions, infectious skin complaints, asthma, back pain, any head injury, and liver disease. Also, if they had ever suffered from the following: neck or major head injury, diabetes, epilepsy, any middle ear diseases, or meningitis. Participants were not allowed to continue if they answered yes to any of these questions due to the perceived increase risk of VE sickness.

Participants were warned that they might suffer from the following: headache, blurred vision, dizziness (eyes open and closed), eyestrain, and sickness, and asked to report any symptom to the experimenter. They were also informed that they could withdraw from the experiment at any time, and they did not have to give a reason for doing so; participants therefore gave informed consent. After the experiment participants were asked to indicate on a scale of 1 (not at all) to 5 (severely), at 15 minute intervals, whether they felt any symptoms.

In the brick tiled/normal FOV condition, two participants suffered a range of symptoms and did not finish the four test trials. One participant reported a slight headache and blurred vision after the second test trial, both symptoms were judged to be a maximum of 2 on the severity scale, and had disappeared at the first 15 minute interval. The other participant reported a headache and eyestrain, both were judged to be a maximum of 4 on the severity scale; dizziness (eyes open and closed) and nausea, judged to be a maximum of 3 on the severity scale. All symptoms, however, were reported to have disappeared after 20 minutes. One other participant reported a slight headache and nausea after the fourth test trial, both were judged to reach a maximum of 2 on the severity scale, and disappeared after 15 minutes. In the other three conditions, no participant reported any symptoms of VE sickness or discomfort.

8.4 Visual Fidelity and Physical Movement

In both the desktop conditions (high and low visual fidelity), and the tethered HMD condition, no participant reported any symptoms of VE sickness. All participants finished the experiment and no one withdrew.

In the roving HMD/high visual fidelity scene condition one participant reported suffering from eyestrain, dizziness (eyes open and closed), and nausea, all reaching a maximum of 2 on the severity scale. The participant reported that after 60 minutes all symptoms had fallen to below 2, but more than 1, on the severity scale. One other participant suffered from blurred vision, maximum 2 on the severity scale, but symptoms disappeared after 15 minutes. Both participants completed the experiment and no other participant reported any symptoms of VE sickness.

In the roving HMD/low fidelity visual scene condition four participants reported a range of symptoms: one reported slight eyestrain, another reported dizziness (eyes closed), the third blurred vision and nausea, while the fourth reported blurred vision. All four participants finished the experiment and reported that the symptoms had disappeared at the 15 minute interval. The remaining six participants did not report any symptoms of VE sickness.

8.5 Discussion

The VE sickness data from the present VE experiments seems to be contrary to other studies that have reported on VE sickness. In the desktop conditions a relatively high proportion of participants (5.6%) suffered from VE sickness symptoms, in three cases so severe that the experiment was stopped. This was a surprising result, few studies have reported VE sickness from navigating a desktop VEs and desktop VEs are not generally associated with high instances of VE sickness.

In the HMD conditions, 20% suffered from minor symptoms of VE sickness, however, no one withdrew and the symptoms quickly disappeared. Other studies have reported dissimilar results: Ruddle and Jones (2001) reported that out of 24 participants that used a HMD for a similar search task to the one used throughout this thesis, five withdrew due to VE sickness (20.8%). Also, Ruddle (2004) collated VE sickness data from 10 HMD experiments, two of which were from the original study of Ruddle and Jones (2001). Out of 134 participants 9% withdrew due to VE sickness and two were physically sick. And finally, in another study by Cobb et al. (1999), VE sickness symptoms from nine experiments were analysed, including

148 participants, and showed that 5% of participants suffered serious symptoms and withdrew.

Chapter 9

General Discussion

All four experiments reported in this thesis were conducted to investigate participants' navigational performance and behaviour using the same search task. The first was performed in the real world in the University of Leeds sports hall and investigated the effects of degrading real world sensory information by reducing the participants FOV. The second and third were performed in a VE that was visually similar to real world sports hall, and investigated effects of visual fidelity, movement interface, and FOV. Finally, the fourth experiment was performed in a VE that was visually similar to the University of Leeds virtual reality laboratory and investigated effects of visual fidelity and movement interface. Taken as a whole the experiments investigated navigation from two complementary directions, by degrading real world sensory information (reducing the FOV) and increasing VE fidelity in terms of the visual scene, FOV, and the mechanism used for movement.

In the real world experiment (Chapter 4) the majority of the full-view participants searched the environment with a perimeter strategy in a clockwise or anticlockwise direction, while the majority of the restricted-FOV participants adopted a lawnmower strategy. The restriction of participants' FOV did not increase the number of targets and decoys that were visited, compared to the full-view condition, indicating that participants were able to maintain their orientation throughout the task. What the reduction in the normal FOV took away was made up for by the increased reliance on other feedback sources (e.g., vestibular and proprioceptive), and by the use of a compensatory strategy. Even though the lawnmower strategy created a longer search path with more changes of direction, it did not decrease performance in terms of time taken or target/decoys checked. There were no improvements in performance across trials in either condition, indicating that participants were performing at ceiling level throughout the experiment. In both conditions, using either strategy, participants performed the search task with near perfect efficiency and found the task to be trivial.

The second experiment (Chapter 5) was conducted in a VE and contained four conditions that were used to investigate the effects of two fidelities of visual scene characteristic and two implementations of movement interface. Participants performed quickest and visited fewest targets with forward-only movement and a high-fidelity scene, and in this condition participants mean performance approached a real world level. That said, only half of participants' searches were conducted perfectly, contrasting with more than 90% in the real world. As in the real world experiment, the configurations of the travelled path fell into two main categories, perimeter and lawnmower. However, unlike the real world experiment, participants who used the perimeter strategy travelled substantially further than those who used a lawnmower strategy (a reversal of the results obtained in the previous experiment). One explanation for this is that the lawnmower strategy involves a systematic search of the environment with participants' path only influenced to a small degree by the actual positions of the target and decoy boxes. By methodically passing through the entire environment, participants found most of the targets during the primary search. By contrast, a perimeter strategy is an object location-dependant strategy that attempts to create a path joining all of the boxes together in a continuous loop. Participants using the perimeter strategy often missed a target during their primary search, making another search inevitable and increasing the distance travelled. As noted above, most participants who performed the task with a restricted FOV in the real world adopted the safer, lawnmower strategy.

The third experiment (Chapter 6) further investigated the effects of visual scene fidelity on navigation but in combination with either a normal FOV or a wide FOV. Additional metrics, such as the “perfect search” metric, were developed to further analyse participants' performance and behaviour. Participants came significantly closer to conducting a perfect search with the combination of a wide FOV and a high fidelity (sports hall) scene. Participants with a wide FOV spent proportionally less time at the beginning of a trial and stationary during a trial, and proportionally more time travelling and checking boxes, than participants with a normal FOV. It is suggested that participants with a normal FOV needed to spend proportionally more time stationary looking around, than those with the wide FOV, in order to assimilate the scene because less of the VE was visible at any one time. In this study all participants used a view-direction interface, participants therefore tended to stop to plan ahead because looking around while travelling causes participants to veer off course. Unlike previous experiments, where in some conditions one strategy dominated over the other, approximately half of the participants in every group consistently used either a lawnmower or perimeter strategy. Classification of the errors that prevented participants from conducting perfect searches showed that, in the wide FOV/sports hall condition, there were far fewer instances of local and global neglect than in the other three conditions.

The fourth and final experiment (Chapter 7) investigated visual fidelity and movement interface in a main and supplementary experiment. The three conditions of the main experiment, roving HMD, tethered HMD, and desktop, all used a high fidelity visual scene. Participants that physically walked around the VE searched the space as efficiently as participants that performed the same search task in the real world (Chapter 4). Surprisingly, rotational vestibular and proprioceptive feedback, provided by the tethered HMD interface, produced no more performance benefits over those that used the stationary desktop display. In the roving HMD condition a perimeter search strategy dominated, echoing the behaviour of participants in the full-view real world experiment, whereas in the desktop and tethered conditions two thirds of trials were conducted using the safer lawnmower strategy.

The two conditions of the supplementary experiment, roving HMD and desktop, both used a low-fidelity VE. When participants walked around the VE the reduction in visual fidelity had little impact on participants' performance or behaviour, whereas the low fidelity scene did reduce performance for the desktop condition. A summary of the findings is shown in Figure 9.1 that clearly highlights the substantial difference in performance between real world and roving HMD conditions and all the other conditions.

Research into human navigation has led to the identification of several methods humans employ in finding their way around the environment. Tolman (1948) first used the term 'cognitive map' to describe the enduring and allocentric mental representation that navigators use as a reference for movement. More recently, however, researchers have argued that human navigation relies on representations of environments that are dynamic and egocentric. Wang and Spelke (2002) suggest that humans, like other animals, use and integrate three separate methods for navigation:

1. Dynamically updated path integration.
2. Viewpoint-dependent place recognition.
3. A congruence-finding reorientation system based on the geometry of the environment.

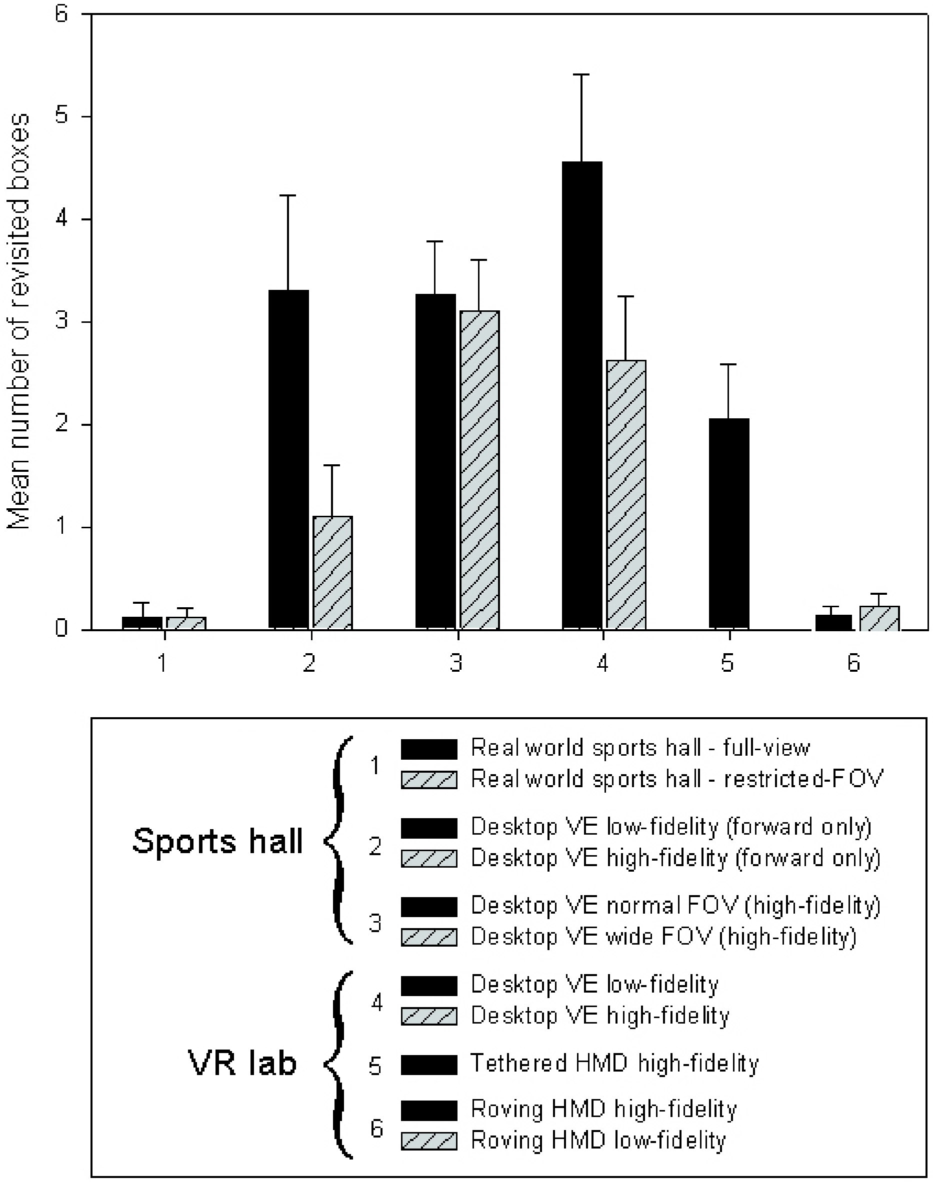


Figure 9.1: Mean number of revisits during a trial for selected conditions within each experiment. Error bars indicate the standard error.

Path integration is the process of using internal information generated from body movement (vestibular and proprioception) and/or external information from optic flow for navigation. Distance and direction information is integrated from velocity and acceleration information into a dynamic egocentric mental representation of the travelled path (Riecke, van Veen & Bülthoff, 2002). Viewpoint-dependent place recognition relies on recognising views by the identification of landmarks: perhaps the most important wayfinding cue. There is growing evidence that even after experiencing multiple perspectives, participants maintain viewpoint-dependent memories (egocentric mental representations) of significant places (Diwadkar & McNamara, 1997). Finally, a congruence-finding system is used to reorient, when path integration is disrupted, by analysing the geometry, but not other properties such as colour, of the surrounding environment (Hermer & Spelke, 1996).

Research into human navigation has often tested participants during simple subtasks of navigation in an attempt to find out innate human abilities and underlying cognitive processes. Such experiments typically investigate the accuracy of participants ability to turn to previously seen objects, or the ability to visually estimate egocentric distances (between self and object) and exocentric distances (between two objects). In contrast, other research has required participants to conduct more complex navigation tasks, such as finding target objects or locations in multispaced (large-scale) environments.

The experiment used during this PhD aimed to investigate navigable ability within virtual spaces by investigating the middle ground between those two positions. Participants were present in the same space as the target objects (i.e. within a small-scale space), and were required to perform translations and rotations while conducting a brief and simple search task.

The advantage of using VEs to research navigation ability is that researchers can often more easly manipulate task conditions i.e. environmental and interface fidelity (Waller, Hunt & Knapp, 1998). The information gained can then not only be of use to VE application designers to discover the affects of system attributes, but can also inform researchers into the cognitive processes used to navigate real and virtual spaces.

Individual differences were not explicitly controlled in the experiments reported in this thesis, thus the studies have certain limitations. Individual differences have been shown to affect performance during spatial tasks in other studies, for example, gender differences have been found in mental rotation tasks, with males performing quicker than females (Linn & Petersen, 1985). Gender has also been found to affect differences in navigational strategies; when participants navigated through a virtual water maze to a target location females relied mainly on landmark information, while males used both landmark and geometric information (Sandstrom, Kaufman & Huettel, 1998). However, other research shows that differences in spatial performance can also be significantly affected by socio-cultural expectations (Sharps, Price & Williams, 1994). Age has also been shown to affect navigation tasks. In a study by Moffat, Zonderman, and Resnick (2001) participants were asked to navigate a virtual maze; older participants took longer to complete each trials, travelled a longer distance, and made significantly more spatial memory errors than younger participants.

The studies reported here also did not consider level 3 metrics concerning participants' decision making rationale (Ruddle & Lessels, 2004), which could have revealed individual differences in cognitive styles and strategies (Chen, Czerwinski & Macredie, 2000). Future studies might address these limitations by controlling for gender, age, cognitive style, or other individual differences that might affect performance and/or behaviour in virtual environment navigation tasks.

The knowledge gained from the studies reported in this thesis could have immediate use for present day VE applications. Designers of VE applications (see Chapter 1) could benefit from the new VE travel taxonomy by offering a way to view and relate methods and devices for travel within a multidimensional design space. This could create a new perspective on the design problem and a new incite to possible application design solutions.

The methods and techniques of analyses of experiment data reported here could be of interest to those responsible for the evaluation of VE applications. The three levels of metric proposed by Ruddle and Lessels (2004) provides a useful categorisation for task analysis, and could be useful to those wishing to understand the underlying causes of poor navigational performance, but could also be used in many other areas of task analysis.

VE applications designers could use the knowledge gained to increase their awareness of how FOV, visual fidelity, and movement interface, effect navigational behaviour and performance. Applications, which are designed to familiarise users for tasks in real world spaces, could use affordable desktop VEs if adequate training is provided, although even with extended training performance is unlikely to be at real world levels all of the time for all users. However, where is it critical that users perform in the same manner as they do in the real world, from the outset, then the results from these studies show the benefits of providing both vestibular and proprioceptive feedback.

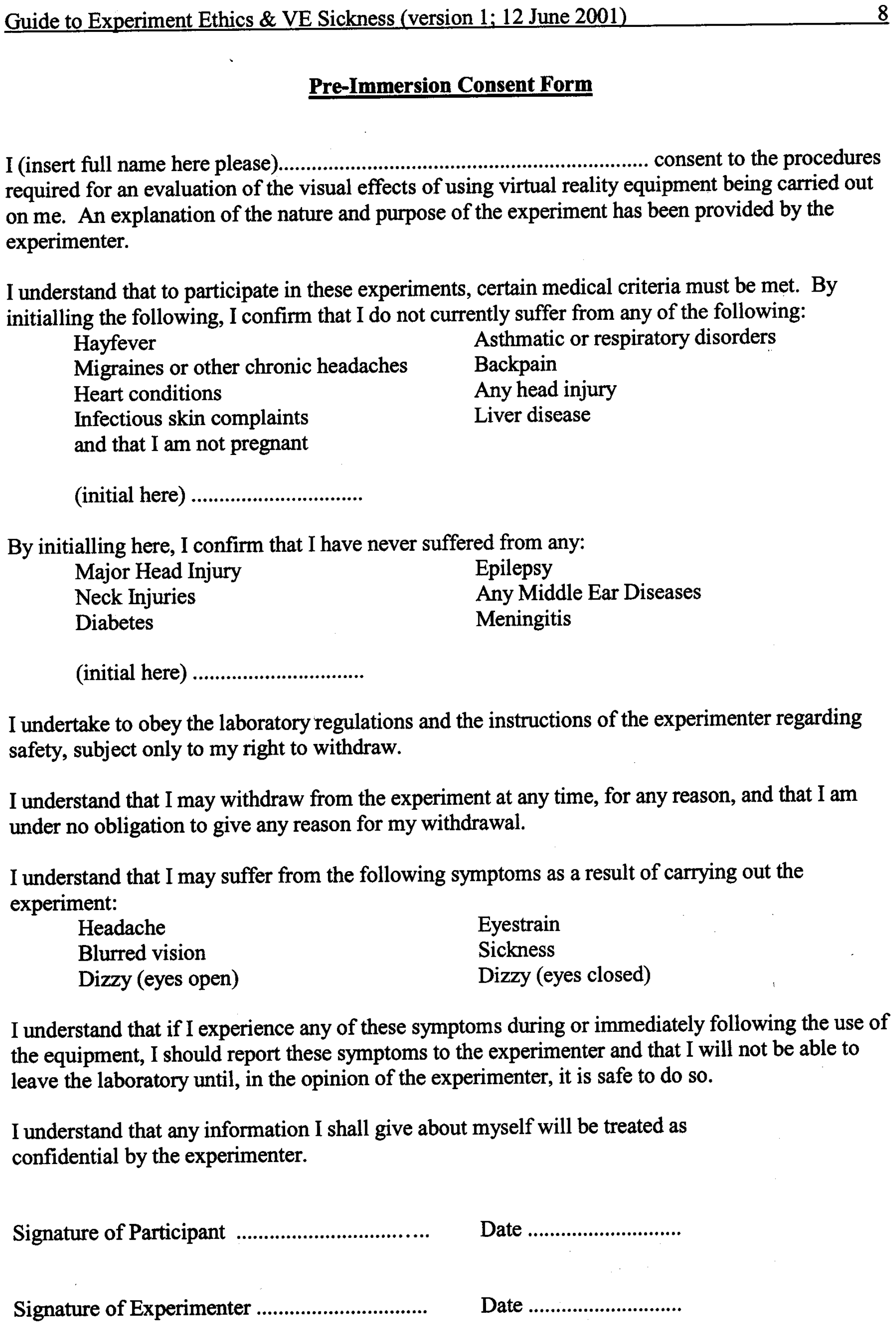
One of the major barriers for the use of VEs for learning spatial information is that users frequently become disorientated and seem to have little idea of where they have travelled. A VE rich in orientation cues theoretically contains sufficient information for participants to keep track of their direction and orientation, but in practice they have difficulty in doing so. An interface that provides for full physical movement increases knowledge of self position and orientation and allows VE navigation performance to rival that of real world navigation. Further research is needed to determine whether these finding extend to larger spaces such as virtual buildings and cities.

In conclusion, the studies presented in this thesis have some important implications for the design of VEs for navigation. First, when the visual fidelity of a VE is increased and is used in combination with either a simple interface or a wide FOV, participants' navigational ability also increases. However, even with this combination of factors, performance remains substantially below that which occurs in the real world. It should be noted that visual fidelity does not necessarily mean realism, rather, the amount of relevant orientation cues and visual information distributed within the scene.

Second, the rotational vestibular and proprioceptive information provided by a tethered HMD does not improve small-scale search performance beyond that which is seen with a stationary desktop display. Third, an interface that provides full (i.e. rotational and translational) vestibular and proprioceptive feedback dramatically improves participants' search performance, to that seen in the real world. Finally, when full proprioceptive and vestibular information is available the fidelity of the visual scene has little impact on performance, demonstrating the dominance of information gained from physical movement for efficient navigation in room-sized spaces.

Chapter 10

Appendix



Guide to Experiment Ethics & VE Sickness (version 1; 12 June 2001) 8

Pre-Immersion Consent Form

I (insert full name here please).. consent to the procedures required for an evaluation of the visual effects of using virtual reality equipment being carried out on me. An explanation of the nature and purpose of the experiment has been provided by the experimenter.

I understand that to participate in these experiments, certain medical criteria must be met. By initialling the following, I confirm that I do not currently suffer from any of the following:

Hayfever
Migraines or other chronic headaches
Heart conditions
Infectious skin complaints
Asthmatic or respiratory disorders
Backpain
Any head injury
Liver disease

and that I am not pregnant

(initial here) ....

By initialling here, I confirm that I have never suffered from any:

Major Head Injury
Neck Injuries
Diabetes
Epilepsy
Any Middle Ear Diseases
Meningitis

(initial here) ....

I undertake to obey the laboratory regulations and the instructions of the experimenter regarding safety, subject only to my right to withdraw.

I understand that I may withdraw from the experiment at any time, for any reason, and that I am under no obligation to give any reason for my withdrawal.

I understand that I may suffer from the following symptoms as a result of carrying out the experiment:

Headache
Blurred vision
Dizzy (eyes open)
Eyestrain
Sickness
Dizzy (eyes closed)

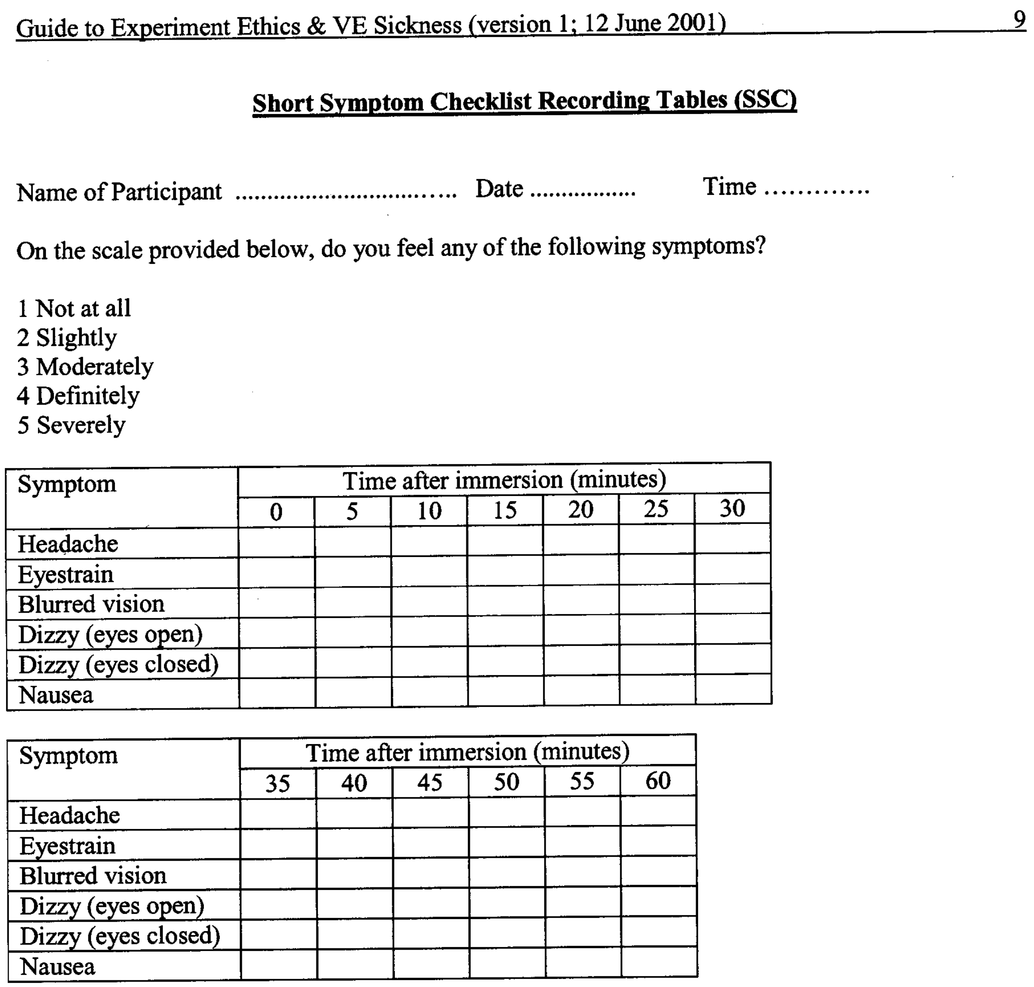
I understand that if I experience any of these symptoms during or immediately following the use of the equipment, I should report these symptoms to the experimenter and that I will not be able to leave the laboratory until, in the opinion of the experimenter, it is safe to do so.

I understand that any information I shall give about myself will be treated as confidential by the experimenter.

Signature of Participant .... Date ....

Signature of Experimenter .... Date ....

Figure 10.1: Page one of the Short Symptoms Checklist



Guide to Experiment Ethics & VE Sickness (version 1; 12 June 2001) 9

Short Symptom Checklist Recording Tables (SSC)

Name of Participant .... Date .... Time ....

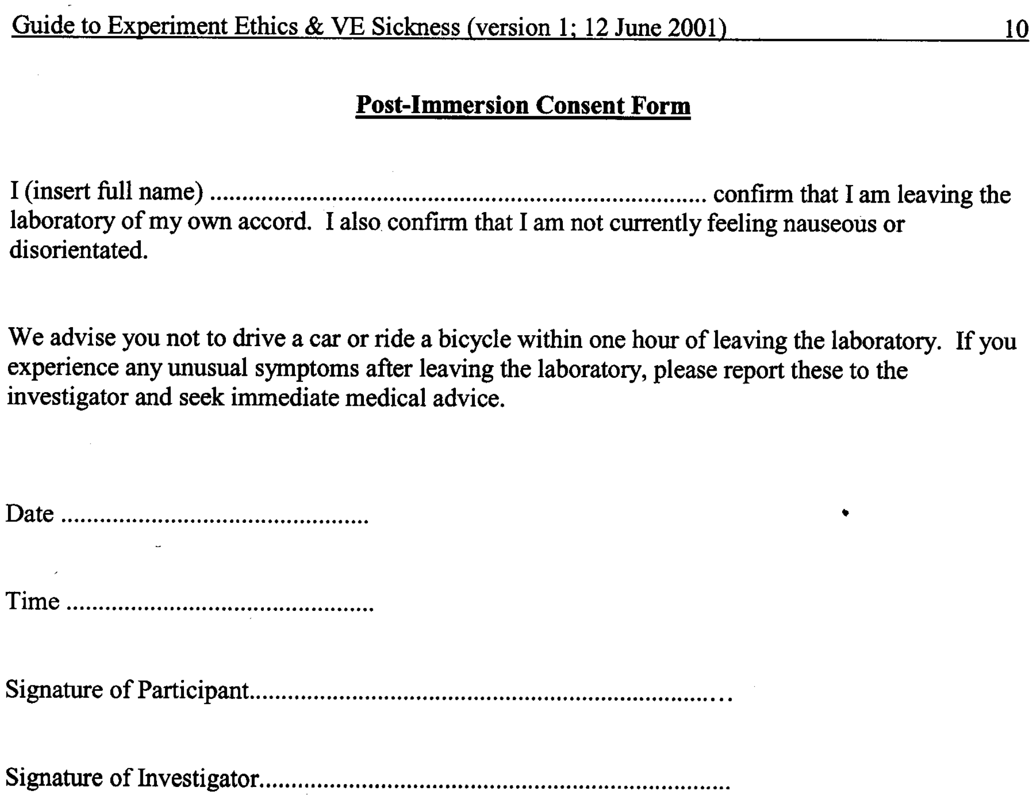
On the scale provided below, do you feel any of the following symptoms?

1 Not at all
2 Slightly
3 Moderately
4 Definitely
5 Severely

Symptom	Time after immersion (minutes)						
	0	5	10	15	20	25	30
Headache							
Eyestrain							
Blurred vision							
Dizzy (eyes open)							
Dizzy (eyes closed)							
Nausea							

Symptom	Time after immersion (minutes)					
	35	40	45	50	55	60
Headache						
Eyestrain						
Blurred vision						
Dizzy (eyes open)						
Dizzy (eyes closed)						
Nausea						

Figure 10.2: Page two of the Short Symptoms Checklist



Guide to Experiment Ethics & VE Sickness (version 1; 12 June 2001) 10

Post-Immersion Consent Form

I (insert full name) .... confirm that I am leaving the laboratory of my own accord. I also confirm that I am not currently feeling nauseous or disorientated.

We advise you not to drive a car or ride a bicycle within one hour of leaving the laboratory. If you experience any unusual symptoms after leaving the laboratory, please report these to the investigator and seek immediate medical advice.

Date ....

Time ....

Signature of Participant..

Signature of Investigator..

Figure 10.3: Page three of the Short Symptoms Checklist

Bibliography

Alfano, P. L. & Michel, G. F. (1990). Restricting the field of view: Perceptual and performance effects. *Perceptual and Motor Skills*, *70(1)*, 35–45.

Allison, R. S., Harris, L. R., Hogue, A., Jasiobedzka, U., Jenkin, H., Jenkin, M., Jaekl, P., Laurence, J., Pentile, G., Redlick, F., Zacher, J. & Zikovitz, D. (2002). Simulating self-motion II: A virtual reality tricycle. In *Virtual Reality, 6* (pp. 86–95). Springer-Verlag.

Bakker, N. H., Werkhoven, P. J. & Passenier, P. O. (1999). The effects of proprioceptive and visual feedback on geographical orientation in virtual environments. *Presence: Teleoperators and Virtual Environments*, *8(1)*, 36–53.

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.

Bolas, M. (1994). Human factors in the design of an immersive display. In *IEEE Computer Graphics and Applications, 14* (pp. 55–59).

Bonakdarian, E., Cremer, J., Kearney, J. & Willemsen, P. (1998). Generation of ambient traffic for real-time driving simulation. *Image Society Conference, Scottsdale, AZ, USA,* (pp. 123–133).

Bowman, D., Kruijff, E., LaViola, J. J. & Poupyrev, I. (2004). *3D User Interfaces. Theory and Practice.* Addison-Westley.

Bowman, D. A., Johnson, D. B. & Hodges, L. (2001). Testbed evaluation of virtual environment interaction techniques. *Presence: Teleoperators and Virtual Environments*, *10(1)*, 75–95.

Bowman, D. A., Koller, D. & Hodges, L. (1997). Travel in immersive virtual environments: An evaluation of viewpoint motion control techniques. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 45–52).

Bruzzone, A., Brandolini, M. & Viazzo, S. (2004). Massive training based on virtual reality equipment applied to logistics and heavy haul trucking. *Summer Computer Simulation Conference, San Jose, California, USA* (pp. 487–492).

Buxton, W. (1983). Lexical and pragmatic considerations of input structures. In *Computer Graphics, 17(1)* (pp. 31–37).

Card, S., Mackinlay, J. & Robertson, G. (1991). A morphological analysis of the design space of input devices. *ACM Transactions on Information Systems*, *9(2)*, 99–122.

Carraro, G. U., Cortes, M., Edmark, J. T. & Ensor, J. R. (1998). The peloton bicycling simulator. In *Proceedings of the Third Symposium on Virtual Reality Modeling Language* (pp. 63–70).

Chance, S. S., Gaunet, F., Beall, A. C. & Loomis, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators and Virtual Environments*, *7(2)*, 168–178.

Chen, C., Czerwinski, M. & Macredie, R. (2000). Individual differences in virtual environments-introduction and overview. *Journal of the American Society for Information Science*, *51(6)*, 499–507.

Chen, J. X., da V. Lobo, N., Hughes, C. E. & Moshell, J. M. (1997). Real-time fluid simulation in a dynamic virtual environment. *IEEE Computer Graphics and Applications*, *17(3)*, 52–61.

Cobb, S. V. G., Nichols, S., Ramsey, A. & Wilson, J. R. (1999). Virtual reality-induced symptoms and effects (VRISE). *Presence: Teleoperators and Virtual Environments*, *8(2)*, 169–186.

Connell, M. & Tullberg, O. (2000). A framework for the interactive investigation of finite element simulations in virtual environments. In *The Second International Conference on Engineering Computational Technology, Belgium.* (pp. 23–28).

Creem-Regehr, S. H., Willemsen, P., Gooch, A. A. & Thompson, W. B. (2003). The effects of restricted viewing conditions on egocentric distance judgments. *Journal of Vision*, *3(9)*, 16a.

Cruz-Neira, C., Sandin, D. & DeFanti, T. (1993). Surround-screen projection-based virtual reality: The design and implementation of the CAVE. In *SIGGRAPH Computer Graphics Proceedings* (pp. 135–142).

Czerwinski, M., Tan, D. S. & Robertson, G. G. (2002). Women take a wider view. In *Proceedings of CHI 2002* (pp. 195–202).

Darken, R. P. & Banker, W. P. (1998). Navigating in natural environments: A virtual environment training transfer study. In *Proceedings of the Virtual Reality Annual International Symposium* (pp. 12–19).

Darken, R. P., Cockayne, W. R. & Carmein, D. (1997). The omni-directional treadmill: A locomotion device for virtual worlds. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology* (pp. 213–221).

Darken, R. P. & Sibert, J. L. (1996). Wayfinding strategies and behaviors in large virtual worlds. In *Conference Proceedings on Human Factors in Computing Systems, CHI 1996* (pp. 142–149).

Deutsch, J. E., Latonio, J., Burdea, G. C. & Boian, R. (2001). Post-stroke rehabilitation with the rutgers ankle system: A case study. *Presence: Teleoperators and Virtual Environments*, *10(4)*, 416–430.

Diwadkar, V. A. & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science*, *8*, 302–307.

Ellis, S. R. (1993). *Pictorial Communications in Virtual and Real Environments*, chapter 2, (pp. 22–40). Taylor & Francis.

Gamberini, L., Cottone, P., Spagnolli, A., Varotto, D. & Mantovani, G. (2003). Responding to a fire emergency in a virtual environment: Different patterns of action for different ssituations. *Ergonomics*, *46(8)*, 842–858.

Grant, S. C. & Magee, L. E. (1998). Contributions of proprioception to navigation in virtual environments. *Human Factors*, *40(3)*, 489–497.

Grantcharov, T. P., Kristiansen, V. B., Bendix, J., Bardram, L., Rosenberg, J. & Funch-Jensen, P. (2004). Randomized clinical trial of virtual reality simulation for laparoscopic skills training. *Britiish Journal of Surgery*, *91(2)*, 146–150.

Harris, L. R., Jenkin, M., Zikovtiz, D., Redlick, F., Jaekl, P., Jasiobedzka, U., Jenkin, H. & Allison, R. S. (2002). Simulating self motion I: cues for the perception of motion. In *Virtual Reality, 6* (pp. 75–85). Springer-Verlag.

Herder, J., Wörzberger, R., Twelker, U. & Albertz, S. (2002). Use of virtual environments in the promotion and evaluation of architectural designs. *Journal of the 3D-Forum Society*, *16(4)*, 117–122.

Hermer, L. & Spelke, S. S. (1996). Modularity and development: the case of spatial reorientation. *Cognition*, *61*, 195–232.

Hindmarsh, J., Fraser, M., Heath, C., Benford, S. & Greenhalgh, C. (2000). Object-focused interaction in collaborative virtual environments. *ACM Transactions on Computer-Human Interaction*, *7(4)*, 477–509.

Jacob, R. J. K. & Sibert, L. E. (1992). The perceptual structure of multidimensional input device selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 211–218).

Kearns, M. J., Warren, W. H., Duchon, A. P. & Tarr, M. J. (2002). Path integration from optic flow and body senses in a homing task. *Perception*, *31(3)*, 349–374.

Kerr, S. J., Neale, H. R. & Cobb, S. V. G. (2002). Virtual environments for social skills training: The importance of scaffolding in practice. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies* (pp. 104–110).

Kirschen, M. P., Kahana, M. J., Sekuler, R. & Burack, B. (2000). Optic flow helps humans learn to navigate through synthetic environments. *Perception*, *29(7)*, 801–818.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S. & Golledge, R. G. (1998). Spatial updating of self position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, *9(4)*, 293–298.

Kline, P. B. & Witmer, B. G. (1996). Distance perception in virtual environments: Effects of field of view and surface texture at near distances. In *Human Factors and Ergonomics Society 40th Annual Meeting* (pp. 1112–1116).

Linn, M. C. & Petersen, A. C. (1985). Emergence and characterisation of gender differences in spatial abilities: A meta-analysis. *Child Development*, *56*, 1479–1498.

Loftin, R. B. & Kenney, P. (1995). Training the hubble space telescope flight team. *Computer Graphics and Applications, IEEE, 15(5)*, 31–37.

May, J. G. & Badcock, D. R. (2002). *Handbook of Virtual Environments, Design, Implementation, and Applications* (pp. 29–63). Lawrence Erlbaum Associates.

Moffat, S. D., Zonderman, A. B. & Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. *Neurobiology of Aging*, *22*, 787–796.

Mohler, B. J., Thompson, W. B., Creem-Regehr, S., Pick, H. L., Warren, W., Rieser, J. J. & Willemsen, P. (2004). Visual motion influences locomotion in a treadmill virtual environment. In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization* (pp. 19–22).

Molet, T., Aubel, A., Capin, T., Carion, S., Lee, E., Magnenat-Thalmann, N., Noser, H., Pandzic, I., Sannier, G. & Thalmann, D. (1999). Anyone for tennis? *Presence: Teleoperators and Virtual Environments*, *8(2)*, 140–156.

Murray, C. D., Bowers, J. M., West, A. J., Pettifer, S. & Gibson, S. (2000). Navigation, wayfinding, and place experience within a virtual city. *Presence: Teleoperators and Virtual Environments*, *9(5)*, 435–447.

Oliva, A., Mack, M. L., Shrestha, M. & Peeper, A. (2004). Identifying the perceptual dimensions of visual complexity of scenes. In *Proceeding of the 26th Annual Meeting of the Cognitive Science Society.*

Pertaub, D.-P., Slater, M. & Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence: Teleoperators and Virtual Environments*, *11(1)*, 68–78.

Péruch, P., May, M. & Wartenberg, F. (1997). Homing in virtual environments: Effects of field of view and path layout. *Perception*, *26(3)*, 301–311.

Peterson, B., Wells, M., Furness, T. & Hunt, E. (1998). The effects of the interface on navigation in virtual environments. *Proceedings of Human Factors and Ergonomics Society 1998 Annual Meeting* (pp. 1496–1505).

Poupyrev, I., Weghorst, S., Billinghurst, M. & Ichikawa, T. (1998). Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques. *Computer Graphics Forum*, *17(3)*, 41–52.

Presson, C. C. & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, *23*, 1447–1455.

Psotka, J., Lewis, S. A. & King, D. (1998). Effects of field of view on judgments of self-location: Distortions in distance estimations even when image geometry exactly fits the field of view. *Presence: Teleoperators and Virtual Environments*, *7(4)*, 352–369.

Riecke, B. E., van Veen, H. A. H. C. & Bülthoff, H. H. (2002). Visual homing is possible without landmarks: A path integration study in virtual reality. *Presence: Teleoperators and Virtual Environments*, *11(5)*, 443–473.

Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *15(6)*, 1157–1165.

Roberts, B., Pioch, N., & Ferguson, W. (2000). Verbal coaching during a real-time task. *International Journal of Artificial Intelligence in Education*, *11*, 377–388.

Robertson, G., Czerwinski, M. & Dantzich, M. V. (1997). Immersion in desktop virtual reality. In *Proceeding of the 26th Annual Meeting of the Cognitive Science Society* (pp. 11–19).

Ruddle, R. A. (2004). The effect of environment characteristics and user interaction on levels of virtual environment sickness. In *Proceedings of IEEE Virtual Reality (VR'04)* (pp. 141–148).

Ruddle, R. A. & Jones, D. M. (2001). Movement in cluttered virtual environments. *Presence: Teleoperators and Virtual Environments*, *10(5)*, 511–524.

Ruddle, R. A. & Lessels, S. (2004). Three levels of metric for evaluating wayfinding. In *Proceedings of Virtual Reality Design and Evaluation Workshop, Nottingham.* Available on CD (ISBN 0 85358 123 1).

Ruddle, R. A., Payne, S. & Jones, D. (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.

Ruddle, R. A., Payne, S. J. & Jones, D. M. (1997). Navigating buildings on "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied*, *3(2)*, 143–159.

Ruddle, R. A. & Péruch, P. (2004). Effects of proprioceptive feedback and environmental characteristics on spatial learning in virtual environments. *International Journal of Human-Computer Studies*, *60(3)*, 299–326.

Sandstrom, N. J., Kaufman, J. & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research*, *6*, 351–360.

Schulte-Pelkum, J., Riecke, B. E., von der Heyde, M. & Bülthoff, H. H. (2002). Perceiving and controlling simulated ego-rotations from optic flow: Influence of field of view (FOV) and display devices on ego-motion perception. In *Object Perception and Memory (OPAM) 2002*. Presented at OPAM 2002, Kansas City.

Sharps, M. J., Price, J. L. & Williams, J. (1994). Spatial cognition and gender: Instructional influences on mental image rotation performance. *Psychology of Women Quarterly*, *18*, 413–425.

Sinai, M. J., Krebs, W. K., Darken, R. P., Rowland, J. H. & McCarley, J. S. (1999). Egocentric distance perception in a virtual environment using a perceptual matching task. In *Proceedings of the 43rd Annual Meeting Human Factors and Ergonomics Society*, Volume 43 (pp. 1256–1260).

Slater, M. & Usoh, M. (1993). Simulating peripheral vision in immersive virtual environments. *Computers & Graphics*, *17(6)*, 643–653.

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 Interface, Special Issue on Virtual Reality Software and Technology*, *2(3)*, 201–219.

Soares, L., Nomura, L., Cabral, M., Dulley, L., Guimaraes, M., Lopes, R. & Zuffo, M. (2004). Virtual hang-gliding over Rio de Janeiro. In *Workshop at IEEE VR04*.

Steck, S. D. & Mallot, H. A. (2000). The role of global and local landmarks in virtual environment navigation. *Presence: Teleoperators and Virtual Environments*, *9(1)*, 69–83.

Sullivan, J., Darken, R. P. & McLean, T. (1998). Terrain navigation training for helicopter pilots using a virtual environment. In *Third Annual Symposium on Situational Awareness in the Tactical Air Environment.*

Sutherland, I. E. (1965). The ultimate display. *In Proceedings of IFIPS Congress, New York City, NY*, *2*, 506–508.

Tan, D. S., Gergle, D., Scupelli, P. G. & Pausch, R. (2004). Physically large displays improve path integration in 3d virtual navigation tasks. In *Proceedings of the 2004 Conference on Human Factors in Computing Systems* (pp. 439–446).

Templeman, J. N., Denbrook, P. S. & Sibert, L. E. (1999). Virtual locomotion: Walking in place through virtual environments. *Presence: Teleoperators and Virtual Environments*, *8(6)*, 598–617.

Tlauka, M. & Wilson, P. N. (1994). The effect of landmarks on route-learning in a computer-simulated environment. *Journal of Environmental Psychology*, *14(4)*, 305–313.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *5(4)*, 189–208.

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. *Proceedings of SIGGRAPH 99* (pp. 359–364).

Waller, D. (1999). Factors affecting the perception of interobject distances in virtual environments. *Presence: Teleoperators and Virtual Environments*, *8(6)*, 657–670.

Waller, D. (2000). Individual differences in spatial learning from computer-Simulated environments. *Journal of Experimental Psychology: Applied*, *8*, 307–321.

Waller, D., Hunt, E. & Knapp, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators and Virtual Environments*, *7(2)*, 129–143.

Waller, D., Loomis, J. M. & Haun, D. B. M. (2004). Body-based senses enhance knowledge of directions in large-scale environments. *Psychonomic Bulletin & Review*, *11(1)*, 157–163.

Waller, D., Loomis, J. M. & Steck, S. (2003). Inertial cues do not enhance knowledge of environmental layout. *Psychonomic Bulletin & Review*, *10(4)*, 987–993.

Wang, R. F. & Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences*, *6(9)*, 376–382.

Ware, C., Arthur, K. & Booth, K. S. (1993). Fish tank virtual reality. *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 37–42).

Ware, C. & Osborne, S. (1990). Exploration and virtual camera control in virtual three dimensional environments. In *Proceedings of the 1990 symposium on Interactive 3D graphics* (pp. 175–183).

Weatherford, D. L. (1985). *The Development of Spatial Cognition*. Hillsdale, NJ: Erlbaum.

Welch, R., Blackmon, T. T., Liu, 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.

Wilson, B. H., Mourant, R. R., Li, M. & Xu, W. (1998). A virtual environment for training overhead crane operators: Real-time implementation. *IIE Transactions*, *30(7)*, 589–595.

Witmer, B. G., Bailey, J. H., Knerr, B. W. & Parsons, K. C. (1996). Virtual spaces and real world places: Transfer of route knowledge. *International Journal of Human-Computer Studies*, *45(4)*, 413–428.

Witmer, B. G. & Kline, P. B. (1998). Judging perceived and traversed distance in virtual environments. *Presence: Teleoperators and Virtual Environments*, *7(2)*, 144–167.

Zanbaka, C., Lok, B., Babu, S., Xiao, D., Ulinski, A. & Hodges, L. F. (2004). Effects of travel technique on cognition in virtual environments. In *Proceedings of the IEEE Virtual Reality Conference (VR04)* (pp. 149–156).

Zheng, X. S., McConkie, G. W. & Schaeffer, B. (2003). Navigational control effect on representing virtual environments. In *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting.* http://www.isl.uiuc.edu/Publications/HFES2003VRPaper-final.pdf.